

Query Match	70.8%;	Score 3595.8;	DB 6;	Length 3715;	
Best Local Similarity	98.6%;	Pred. No. 0;			
Matches 3685;	Conservative	0;	Mismatches	7;	Indels 45; Gaps 4;
QY	1335	tgacgcgagagaaacctctgtttttcccccactctctctccacctctctctgccttccc	1394		
DB	4	TTCCGGCGGAGAGAACCTCTGTTTTCCTCCACTCTCTCCACCTCCCTTCCTTCCC	63		
QY	1395	caccgcagtcgagcgacagatcaaaagatgaaaaagcagtcaggtcttcagtagcca	1454		
DB	64	CACCCCGAGTCGGAG-CAGAGATCAAAAGATGAAAGGCAGTCAGGTCTTCAGTAGCCA	122		
QY	1455	aaaaaacaacaaacaaacaaacaaacaaacaaacaaacaaacaaacaaacaaac	1514		
DB	123	AAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA	182		
QY	1515	cttattgaacctactcagtgagacactgaatttggaaagtgagagattttgttttttc	1574		
DB	183	CTTATTGTGCACTTACTTCAAGTGACACTGAATTTGGAAGTGGAGATTTTGT	242		
QY	1575	ttttaagatctggcactctttgaactaccctcaagtattaaagacagactgtgac	1634		
DB	243	TTTTAGATCTGGCAATCTTTTGAATCTACCTTCAAGTATTAAAGACAGACTGTGAC	302		
QY	1635	ctagcgggagatcttgtccacgtgtgtctctctctcagcagactttgaggtgtca	1694		
DB	303	CTAGCAGGCAGATCTGTCCACCGTGTCTCTCTCTGCACGAGACTTTGAGGCTGTCA	362		
QY	1695	gagcgtttttgaggttgctcccgcaagtttctctctgagacttcccacaggtggg	1754		
DB	363	GAGCGCTTTTGGGTGGTGTCTCCGCAAGTTTCTCTCTGAGGCTTCCCGCAGGTGGG	422		
QY	1755	cagctaactcagcactaccgcatcatcacagcctgttgaaactctctgacaagagaa	1814		
DB	423	CAGTAGCTCAGGGACTACCGCATCATCACAGCCTTTGAACTCTTCAGCAAGAA	482		
QY	1815	gggaggggggtaagggaagtaggtggaagattcagccaagctcaagatggaagtga	1874		
DB	483	GGGAGGCGGGTAAGGGAAGTAGTAGTGGGAAGATTACAGCAAGCTCAAGATGGAAGTGA	542		
QY	1875	gttaggctgggaaggtctaccctcgccgcgcgtccagacactaccgagagctttcca	1934		
DB	543	GTTAGGGCTGGGAAGGTGTACCTCGCGCGCGTCCAGACCTACCGAGGAGCTTTTCCA	602		
QY	1935	gaactgttcagagcgtgcggaagtgaaccagaccgggcccagagcaccagagc	1994		
DB	603	GAATCTGTTCCAGAGCGTGCGCAAGTAGTCCAGAACCCGGGCCCCAGCACCCAGAGGC	662		
QY	1995	cgagcgcagcaactcccgccagcttgcgtgcgtgcagcagcagcagcagcagca	2054		
DB	663	CGCAGCGCAGCACTCCCGCGCCAGTTTGTGCT-----	701		
QY	2055	gcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	2114		
DB	702	---GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	758		
QY	2115	tagcccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag	2174		
DB	759	TAGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	818		
QY	2175	agggccacaggctactcgtcgtgatgaggaaacgaacacttcaagccagctcgcc	2234		
DB	819	AGGCCCCACAGGCTACCTGTCTGTGATGAGAACACAACTTACAGCCCGCAGTCGCG	878		
QY	2235	cctgagtgccacccgagagaggttgcgtccagagcctggagccgcgtggccgcag	2294		
DB	879	CCTGGAGTGCACCCCGAGAGAGGTTCGCTCCACAGACCTTGAGCCCGCGTGGCCGCCAG	938		
QY	2295	caaggggtcgcgagcagctgcagcactccggacagagtagtactcagctgccccatc	2354		
DB	939	CAAGGGGCTCGCAGCAGCTGCCAGCACTCCGGAGGAGTAGTACTAGCTGCCCATC	998		

```

|||||
Db 2079 TTGTCTAAAAAGCGAAATGGCCCTCGATGGATAGTACTCTCGGACCTTACGGGCACAT 2138
QY 3486 gcgtttgagactgcagggaacatgttttgccattgactatactttccacccagaa 3545
Db 2139 GCgTtTGGAGAcTgCCAGGACcATgTtTtTGGCCATtGAcTATtACTtTtCCACCCAGAA 2198
QY 3546 gacctgcctgatctgtgagatgaagcttctgggtgtcactatggagctctcacatgtg 3605
Db 2199 GAcCTGcCTGATcTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCAcATGTGG 2258
QY 3606 aagctgaaggttcttctcaaaagagccgctgaagggaacagaagtaacctgtgcgccag 3665
Db 2259 AAGCTGAAGGTCTTCTCAAAAGAGCGCTGAAGGGAACAGAACTACTGTGCGCCAG 2318
QY 3666 cagaaatgattgcactatgataaattccgaagaaataattgccatcttgtctgtcg 3725
Db 2319 CAGAAATGATTGCACTATTGATAAATTCGGAAGGAATAATTFCCATCTGTGCTCTCG 2378
QY 3726 gaaatgttatgaagcagggtgactctggagccgggaagctgaagaaacttggtaact 3785
Db 2379 GAAATGTTATGAAGCAGGGATGACTCTGGAGCCCGAAGCTGAAGAACTTGGTAATCT 2438
QY 3786 gaaactacagggaaggagaggtcttcagcaccacacagcccaactgaggaacaccca 3845
Db 2439 GAAACTACAGGAGGAAGAGAGGCTTCCAGCACACCACCCACTGAGGAGACAACCCA 2498
QY 3846 gaagctgaacagtctcacatgaaggctatgaatgcagcaccacagcccaacttctcgaatgct 3905
Db 2499 GAAGCTGACAGTGTCAcATtTGAAGGCTATGAATGTCAAGCCCATCTTCTGAATGTCT 2558
QY 3906 ggaagccattgagccagggtgtagtgtgtgtgacagacacaacacagccagactcctt 3965
Db 2559 GGAAGCCATTGACCAGGTGTAGTGTGTCTGGACAGACACAACACCCAGCTCCCT 2618
QY 3966 tgcagctgtctctctagcctcaatgaactgggagagacagcttgtcacagtggtcaca 4025
Db 2619 TGCAGGCTTTGCTCTCTAGGCTCAATGAACCTGGGAGAGACAGCTGTACAGTGGTCAA 2678
QY 4026 gtgggccaagccttgccctgcctccgaacttacaactgagcagcagatggctgtcat 4085
Db 2679 GTGGGCCAAGCCCTGCCCTGCCGAACTTACACGTGGAGACAGCTGTACAGTGGTCA 2738
QY 4086 tcagtactcctggatgggctcatggtgtgtgccatgggctggcagctcttcacaaagt 4145
Db 2739 TCAGTACTCTCGATGGGCTCATGTGTGTGCCATGGCTGGCGATCCTTCACCAATGT 2798
QY 4146 caactccagatgctctacttcgcccctgatctgtgtttcaatgagtaocgcgcatgcaca 4205
Db 2799 CAACCTCCAGGATGCTCTACTTCGCCCTGATCTGGTGTTCATATGATACCGCATGCACA 2858
QY 4206 gtcccgatgtacagccagtgctgcgaatgagggcacctctctcaagagtttggatgct 4265
Db 2859 GTCCCGATGTACAGCCAGTGTCCGAATGAGGACCTCTCTCAAGAGTTGGATGGCT 2918
QY 4266 ccaaatcaccocccaggaattcctgtgcatgaaagcaactgtcactcttcagcattatcc 4325
Db 2919 CCAAAATCACCCCCAGGAATTCCTGTGCATGAAGCACTGTCTTTCAGCATTTATCC 2978
QY 4326 agtggatggctgaaaaatcaaaaattctttgatgaacttcgaatgaactacatacaga 4385
Db 2979 AGTGGATGGCTGAAAAATCAAAAATTCCTTGTGAACCTTCGAATGAACCTACATCAAGGA 3038
QY 4386 actcgatcatattgcatgcaaaagaaaaatcccatcctctcagaagcttcta 4445
Db 3039 ACTCGATCGTATCATTCATGATGAAGAAAAAATCCACATCTCTGCTCAAGACGCTTCTA 3098
QY 4446 ccagctcacaagctcctggactccgctcagcctattgcgagagagctgcacatgctcac 4505
Db 3099 CCAGCTCAACAAGCTCCTGGACTCGGAGCTATTTCGAGAGAGAGCTGCATCAGTTTAC 3158
QY 4506 ttttgacctgctaataagtcacacatggtgagcgtgaccttccggaaatgatggcaga 4565
|||||
```

```

Db 3159 TTTTGACCTGCTAATCAAGTCACACATGCTGAGCTGGACTTTTCGGAAATGATGGCAGA 3218
QY 4566 gatcatctctgtgcgaagtgcacagatccttctctgggaaagtcaagcccatctatttcca 4625
Db 3219 GATCATCTCTGTGCAAGTGCCCAAGATCCTTCTGGGAAAGTCAAGCCCATCTATTTCGA 3278
QY 4626 caccagtggaagcattgaaaccctatttcccccacccagctcatgcccccttccagatg 4685
Db 3279 CACCAGTGAAGCAFTGGAAACCCCTATTTCGCCACCCAGCTCATGCCCCCTTTTCAGATG 3338
QY 4686 tcttctgcctgttataactctgcactactcctctgcagctgccttggggaaatttctctat 4745
Db 3339 TCTTCTGCTGTATAACTCTGCACACTCTCTCTGCAGTGGCTTGTTTAAATTCCTCTAT 3398
QY 4746 tgatgtacagctgtcattgaacatgttctcctgaattcttatttgcctgggctttttttctc 4805
Db 3399 TGATGTACAGTCTGTCAATG-----GAATCTATTTCCTGGGCTTTTTCCTTC 3447
QY 4806 ttctctccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctc 4865
Db 3448 TTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3507
QY 4866 ttgtcttcccatgtggcctctctctgtgtttgaaatggtgtgtatgcctttaaactctg 4925
Db 3508 TTTGCTTCCCATTTGGCTCTCTATCTGTGTTTGAATGGTGTGTATGCTTTTAAATCTG 3567
QY 4926 tgatgtactcctcatgtggccagtgctcaagtgtgcttgttttacagcactactctgtgcca 4985
Db 3568 TGATGATCCTCATATGGCCAGTGTCAAGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 3627
QY 4986 gccacacaaagtttacttattctattctatgccaggggaagtttagagagtaagattatctg 5045
Db 3628 GCCACACAAAGCTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGG 3687
QY 5046 ggaaatcaaaacaaaaa 5062
Db 3688 GGAAATCAAAACAAAAA 3704

RESULT 2
LOCUS HUMARB 3715 bp mRNA PRI 31-OCT-1994
DEFINITION Human androgen receptor mRNA, complete cds.
ACCESSION M3263 N18624
VERSION M3263.1 GI:178893
KEYWORDS androgen receptor; dihydrotestosterone receptor.
SOURCE Human prostate, cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3715)
AUTHORS Chang,C.S., Kokontis,J. and Liao,S.T.
TITLE Structural analysis of complementary DNA and amino acid sequences
of human and rat androgen receptors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (19), 7211-7215 (1988)
MEDLINE 89017168
REFERENCE 2 (sites)
AUTHORS Chang,C.S., Kokontis,J. and Liao,S.T.
TITLE Molecular cloning of human and rat complementary DNA encoding
androgen receptors
JOURNAL Science 240 (4850), 324-326 (1988)
MEDLINE 88178111
COMMENT Draft entry and computer-readable sequence for [1] kindly provided
by S.Liao, 01-MAR-1989.
Four proteins of different sizes were found. These proteins could
be explained if alternative start codons at positions 1084, 1252
and 1315 were used.
FEATURES
Location/Qualifiers
1..3715
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="Xq11.2-q12"
532..3288
gene
```


Query Match		69.8%;	Score 3547;	DB 6;	Length 3569;		
Best Local Similarity		99.7%;	Pred. No. 0;				
Matches 3569;		Conservative	0;	Mismatches	0;	Indels	12;
</							

|||||
Db 1969 ACATCGCTTTGGAGACTGCCAGGACCAATGTTTGGCCATTGACTATTACTTTCACCCC 2028
Qy 3542 agaagacctgctgatctgtggagatgaagctctctgggtgctcaatgagctctcact 3601
Db 2029 AGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACAT 2088
Qy 3602 gtggagctgcaaggtcttcttcaaaagagccgctgaagggaacagaagtaacctgtgag 3661
Db 2089 GTGSAAGCTGCAAGGCTCTTCTCAAAAGAGCCGCTGGAAGGAAACAGAAAGTACCTGTGCG 2148
Qy 3662 ccagcagaaatgatgctactattgataattccgaaggaataattgtccactctgtgtgc 3721
Db 2149 CCAGCAGAAATGATTCACACTATTGATAAATCCGAAGGAATAATGTCCATCTTGTGCTGC 2208
Qy 3722 ttccgaaatgttatgaagcaggatgaactctgaggccggagccggaagctgaagaaacttgta 3781
Db 2209 TTCGGAATGTTATGAAGCAGGGATGACTCTGGGAGCCGGAAAGCTGAAGAACTGTGTA 2268
Qy 3782 atctgaaactacaggaaggaagaggttccagcaccaccagcccaactgagagacaa 3841
Db 2269 ATCTGAAACTACAGGAGGAAGAGAGGGCTTCAGCACCACCAGCCCACTGAGGAGACAA 2328
Qy 3842 ccagaagctgacagtgccacacattgaaggctatgaatgtcagccctcttcttgaatg 3901
Db 2329 OCCAGAAGCTGACAGTGTACACACTGAAGGCTATGAATGTGAGCCCATCTTCTTGAATG 2388
Qy 3902 tcttgaagccattagccaggtgtagtgtgtctggaacagcaacacacagcccgact 3961
Db 2389 TCTGGAAGCCATTGACCGAGGTGATGTGCTGGACACACAAACACCCAGCCGACT 2448
Qy 3962 ctttgcagcctgtctctagcccaatgaactggaagagagacagctgttacacgtgg 4021
Db 2449 CFTTGCAGCCTTGTCTCTAGCCTCAATGAACCTGGAGAGAGACAGCTTGTACACGTGG 2508
Qy 4022 tcaagtggcccaaggccttgctgcttcgcccaacttacctggacgaccagatggctg 4081
Db 2509 TCAAGTGGCCCAAGGCTTGCCTGGCTTCGCCAACTTACACGTGGAGCACCAGATGGCTG 2568
Qy 4082 tcattcagttactcctggatggggctcattggttggccatgggctgggagatccttcaaca 4141
Db 2569 TCATTTCAGTACTCCTCGATGGGCTCATGGTGTGTGCCATGGGCTGGCGATCCTTTCACCA 2628
Qy 4142 atgtcaactcagagatgctctacttccgcccctgactctgttttcaatgtagtaccgcatgc 4201
Db 2629 ATGTCAACTCCAGGATGCTCTCTCGCCCTGATCTGCTTTTCAATGAGTACCCGATGC 2688
Qy 4202 acaagtcgcggatgtacagccagtggtccgaatgaggcaccctctctcaagagtttggat 4261
Db 2689 ACAAGTCCCGGATGTACAGCCAGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGAT 2748
Qy 4262 ggctccaaatcaacccccggaaattcctgtgcatgaaagcaactgctactcttcagcatta 4321
Db 2749 GGCTCCAAATCACCCCCAGGAATTCCTGTGATGAAGCACTGTCTACTCTTCAGCATTTA 2808
Qy 4322 ttccagtggtatggctgaaatacaaaaattcttggatgaacttccgaatgaactacatca 4381
Db 2809 TTCCAGTGGATGGCTGAAATAATCAAAATTCCTTGTATGAATTCGAATGAACTACATCA 2868
Qy 4382 aggaactcgatcgatcattgtgatgcgaagaaataatccacatcctgctcaagacgt 4441
Db 2869 AGGAACTCGATCGTATCATTTGCATGCAAAAGAAAAATCCCATCTGCTCAAGACGCT 2928
Qy 4442 tctaceagctcacaagcctcctgactcctgagcctattgcagagagactgcatcagt 4501
Db 2929 TCTACCAGCTCACAAAGCTCCTGGACTCCGTGCGAGCTATTGGCAGAGAGCTGCATCAGT 2988
Qy 4502 tcacttttgacctgactaatcaagtccacatggtgagcgtgggacttccggaatgatgg 4561
Db 2989 TCACCTTTGACCTGCTAATCAAGTCACACATGCTGAGCGTGGACTTCCGGAATGATGG 3048
Qy 4562 cagagatcatctctgtgcaagtgcccaagatcctttctgggaagtaagccatcatt 4621
|||||

Db 3049 CAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGAAAGTCAAGCCCATCTATT 3108
Qy 4622 tccaaacccagtgaaacattgaaacccctatttcccccacccagctcatgcccccttca 4681
Db 3109 TCCACACCCAGTGAAGCATTTGGAAACCCCTATTTCGCCACCCAGCTCATGCCCTTTCA 3168
Qy 4682 gatgtctctgcctgttataaactctgactactcctctcagtgcccttggggaattcct 4741
Db 3169 GATGTCTCTGCGCTGTATAAATCTGCACCTACTCTCTGCAGTGCCTTGGGAATTCCT 3228
Qy 4742 ctattgatgtacagtcgtctgtaaacatgttccgtaattctatttctgctgggctttttt 4801
Db 3229 CTATTGATGATACAGTCTGTATGAACATGTCTGAATTTCTATTTCCTGGGCTTTTTTT 3288
Qy 4802 tctcttc 4861
Db 3289 TCTCTTCT 3348
Qy 4862 agactttgttcccatgtgtgctcctctctctctctctctctctctctctctctctct 4921
Db 3349 AGACTTTGTTCCCATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3408
Qy 4922 tctgtgatgtcctcctatagcccgagtgcaagtgctgctgtgtgtttacagcactactctgt 4981
Db 3409 TCTGTGATGATCTCTCATPATGCGCCAGTGTCAAGTGTGCTTGTGTGATGGTGTGTATG 3468
Qy 4982 gccagccacacaaacgtttacttatcttctgcccagcggaagtttagagagactaagattat 5041
Db 3469 GCCAGCCACACAAAGTCTTACTTATCTTATGCCCAGGGAAGTTTAGAGAGCTAAGATTAT 3528
Qy 5042 ctgggggaatcaaaaacaaacaaacaaacaaacaaacaaacaaacaaacaaacaaac 5082
Db 3529 CTGGGGAATCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 3569

RESULT 6
HSDJ80804/c
LOCUS
DEFINITION
Human DNA sequence from clone RP4-80804 on chromosome Xq11.1-12.
Contains the 5' end of the AR gene for androgen receptor
(dihydrotestosterone receptor), a PABPN1 (poly(A)-binding protein,
nuclear 1) (PABP2) pseudogene, ESTs STSS, GSSs and two putative CpG
islands, complete sequence.
ACCESSION AL049564
VERSION AL049564.10 GI:4902757
KEYWORDS HTG; androgen receptor; AR; CpG island; PABP2; PABPN1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 139033)
AUTHORS Chapman,J
DIRECT SUBMISSION
SUBMITTED (03-APR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT
On May 27, 1999 this sequence version replaced gi:4757056.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WormPEP; Information
on the WormPEP database can be found at

repeat_region /note="LIME1 repeat: matches 4835. .6158 of consensus" 39987. .40530

repeat_region /note="L2 repeat: matches 2130. .2710 of consensus" 40676. .40809

repeat_region /note="L2 repeat: matches 2579. .2710 of consensus" 40945. .41137

misc_feature /note="L2 repeat: matches 1990. .2185 of consensus" complement(41397. .41743)

misc_feature /note="match: GSS: Em:AQ087063" complement(41397. .41704)

misc_feature /note="match: GSS: Em:AQ091315" 41818. .42458

misc_feature /note="match: GSS: Em:AQ357243" 41819. .42063

repeat_region /note="match: GSS: Em:AQ878802" 42045. .42079

repeat_region /note="L2 repeat: matches 2641. .2675 of consensus" 42132. .42421

Query Match 62.7%; Score 3185.4; DB 9; Length 139033;

Best Local Similarity 97.7%; Pred. No. 0;

Matches 3428; Conservative 0; Mismatches 40; Indels 40; Gaps 19;

Qy 1 gagctctggacaaaatgagcgcctatgtacatggcaagtgcttttttagtggttgtg 60

Db 36813 GAGCTCTGGACAAATGAGCGCCTATGTACATGCAAGTGTTTGTAGTGTGTG 36754

Qy 61 tttaacctgtctgtgggtgatttgcctttgagagctcgtgagataatgcattgtaa 120

Db 36753 TTTACCTGCTGTGCTGGTGATTTTGCCTTTGAGAGTCTGGATGAGAAATGCA 36694

Qy 121 aggaattccagacagaaagagagagagagagagagagagagagagagagagag 180

Db 36693 AGGCAATTCAGACAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 36634

Qy 181 ggctgaggggttcctagagcaaatggcacaatgccagagagcccgatctatccctatgacg 240

Db 36633 GGCTGAGGGTTCTAGAGCAATGGCACATGCCACGAGCGCCGATCTATCCCTATGACG 36574

Qy 241 gaactctaaagtttcagactaactctctgtgctgtgctgctgctgctcctcag 300

Db 36573 GAA-TCTAAGGTTTCAG---CAAGTATCTGCTGGCTTGGTCA--TGGCTTGTCTCC 36520

Qy 301 ttgtagagactctccactctccactctcctatgcgcctctatcagctcctgaaagaacccn 360

Db 36519 TTGTAGAGAGACTCTCCACACTCTCCACTCTCCACTCTCCACTCTCTTAACAGTCTG 36460

Qy 361 tggcnagccaggagcnaggtattentatcgtcctctctctctctctctctctngc---ctcacct 416

Db 36459 TGGC-AGCCAGGAGC-AGGTATTCTATCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 36402

Qy 417 ngttgnttttttagattggnottngnaacaaatt---tgtatcgtggcctccaggaaatc 473

Db 36401 GTTGGTTTTTTGTAGATTGGGCTTTGGAACCAAAATTTGGTGTAGTCTGGCTTCCAGGAA 36342

Qy 474 tggag-cctggcgctaaaccttggttttagaaagcagagagctattcaggaagca-ggggt 531

Db 36341 TGGAGCCCTGGCGCTTAACCTTGGTTTAGAAAGCAGAGAGCTATTACAGAGCAGGGGT 36282

Qy 532 cctccagggctagagctagcctctcctgcctgcctgcacg-tgcgcagcacttgtttct 590

Db 36281 CCTCCAGGGCTAGAGCTAGCCTCTCTCTGCGCTCTGCGCCACGCTGCGCCAGCATTGTTCT 36222

Qy 591 ccaagcnactltagcagcgcttagcgcgcgtgagggagagagagagagagagagagag 650

Db 36221 CCAAGGCACATAGCAGCGCGTTAGCGCGGTGAGGGGAGGGGAGAGAAAGAGGGGAG 36162

Qy 651 gggagggaaaagaggttgggaaggaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 710

Db 36161 GGGAGGGAAGAGAGGTGGGAAGCAGAGGAGCGCGCGCGCTGGGGCGGGACCCGACTC 36102

Qy 711 gcannaactgttcattgtctcctccactccagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 770

Db 36101 GCA--AACTGTTGCATTGTCTCTCCAGCTCCAGCGCCCTCCGAGATCCCGGGAGCC 36044

Qy 771 agcttgctggagagcgggaacggtccggagcaagccagagagcagagagagagagagagag 830

Db 36043 AGCTTGCTGGGAGAGCGGG-ACGGTCCGGAGCAAGCCAGAGGAGAGGAGGACAGAG 35985

Qy 831 ggaataaggaagccnagctagcgcctccagtgctgtacagnagccgaa-ggacgcaccacg 889

Db 35984 GGAAGAGGG-CCGAGCTAGCCGCTCCAGTGTGTACAGGAGCGAAGGAGCGACCCACC 35926

Qy 890 ccagcccccagcccggctccagcgacagcnaacgcctcttgca-----ngcgttcgaagc 943

Db 35925 CCAGCCCCAGCCCGCTCCAGCGACAGCAACAGCGCTCTTGCGAGCGCGCGCTTCGAAGC 35866

Qy 944 cgcgcgcgcgcgcgcctctctctcgcgtgaagtcttaaaagcgcctaaagactcgg 1003

Db 35865 CGCGCCCGGAGCTGCGCTTCTCTCGGTGAAGTCTTTAAAGCTGCTAAAGACTCGG 35806

Qy 1004 aggaagcaagaaagtgcctggttaggactgacgctgcttctcctctcctcctcctccacc 1063

Db 35805 AGGAAGCAAGAAAGTGCCTGTTAGGACTGACGGTGCCTTTGTCTCTCTCTCCACC 35746

Qy 1064 cgcctccccccactgcctctcccccctcccccctctctctctcctcctcctcctcctcctc 1123

Db 35745 CGCGCTCCCGCCACCGCTGCTTCCCGCTCTCCCGCTCTTCTCTCCCGAGCTGCCTCAG 35686

Qy 1124 tggctactctcagcaaacccccctcaccacctctcccccacgcgcgcgcgcgcgcgcgcgc 1183

Db 35685 TCGGCTACTCTCAGCAACCCCGCTACACCCCTCTCCCCACCGCGCCCCCGCCCCCG 35626

Qy 1184 tcggcccgagcgnatgncagncgcgagtttcagagaggttaactcccttttgctgcgagcgg 1243

Db 35625 TCGGCCAGCGCTGCCACGCCGAGTTTCAGAGAGGTAACCTCCCTTTGGCTTCGAGCGGG 35566

Qy 1244 cgagntcagctgcacattgcataagagaggtctcttaggag-caggcgactcgggagcgctt 1302

Db 35565 CGAG-CTAGCTGCATTTGCAAGAGGCTCTTAGAGGCCAGGGAGCTTGGGAGCGGCTT 35507

Qy 1303 cagactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1362

Db 35506 CAGCACTCAGCCAGCGCGCGCTGGTGTAGCTGCACCGGAGAGAACCTCTGTTTTC 35447

Qy 1363 cccactctctccaccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1422

Db 35446 CCCACTCTCTCCACCCTCTCTCTCCCTTCCCTCCCGAGTGGGAGCCAGAGATCAA 35387

Qy 1423 agatgaagagcagtcaggtctcagtagcacaacaaacaaacaaacaaacaaacaaacaa 1482

Db 35386 AGATGAAGAGGCGAGTCAGGTCTTCAGTAGCCAAAACAAACAAACAAACAAACAA 35327

Qy 1483 ccgaataaagaaagaaataaactcagttcttatttgacactacttcagtggaact 1542

Db 35326 CCGAAATAAAGAAAGAAAGATAATAACTCAGTCTTATTTGACCTACTCTCAGTGACACT 35267

Qy 1543 gaatttgaagctggaggaatttggtttttttctttaaagatctgggcatcttttgaatct 1602

Db 35266 GAATTTGGAAGTGAGGATTTTGTCTTTTCTTTTAAATCTGGGCATCTTTTGAATCT 35207

Qy 1603 accctcaagtattaaagagacagactgtgagcctagcagggcagatcttgtccaccgtgt 1662

Db 35206 ACCCTTCAAGTATTAAAGACAGACTGTAGCTAGCAGGGCAGATCTTGTCCACCGTGT 35147

Qy 1663 gtctctctcgcacgagactttgagcgtcgaagcgtcttttgcgtgtgtgctcccgca 1722

Db 35146 GTCTTCTCTGCAGAGACTTTGAGGCTGTACAGCGCTTTTGTGCTGTTGCTCCCGCA 35087

Qy 1723 agtttctctctgagacttcccgcaggtgggcagctagctgcagcactaccgcacatcat 1782

Db 35086 AGTTTCTCTCTGAGGCTTCCCAGGTTGGCAGCTAGCTAGCTGACGAGCTACCGCATCAT 35027

Qy 1783 cacagcctgtgaactctctcagcaagaagagagagagagagagagagagagagagagag 1842

Db 35026 CACAGCTCTTGAACCTCTTCTGAGCAAGAGAGGGGAGCGGGGTAAAGGAAGTAGGTGG 34967

```
QY 1843 aagattcagcaagctcaagatgaagtgagtttaggtggaaggttctacccctcgg 1902
Db 34966 AAGATTCAAGCTCAAGATGGAAGTGCAGTTAGGCTGGGAAGGTTCTACCTCGG 34907
QY 1903 ccgcctcacaagactaccagaggtcttcagaatctgttccagagcgtgctgcaagt 1962
Db 34906 CCGCGCTCAAGAGCTTACCGAGAGGCTTTCCAGAACTGTGTCCAGAGCGTCCGCGAAGTG 34847
QY 1963 atcagaacccggcccccagacacccagagcccgagcagcagcaccctccggccaggt 2022
Db 34846 ATCCAGAACCCGGGCCCCAGGACCCAGAGCCCGAGGCGAGCACCTCCCGGCGCCAGT 34787
QY 2023 ttgtctgtctgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2082
Db 34786 TTGCTGTCTGCT-----GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 34733
QY 2083 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2142
Db 34732 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 34673
QY 2143 ggtgaggtgttctccccaaagcccatcgtagagggcccccacaggtacctggtcctgagt 2202
Db 34672 GGTGAGGATGTTCTCCCAAGCCCCATCGTAGAGGCCCCACAGGCTACCTTGGTCTTGAT 34613
QY 2203 gaggaacagcaaccttcacagccgagctgagccctgagtgccacccccagagaggttgc 2262
Db 34612 CAGGAACAGCAAACTTCACAGCCCGACGTGGCCCTGGAGTGCCACCCGAGAGAGGTTGC 34553
QY 2263 gtccagagcctgagcgcgcgtgagccgacagcaagaggggtgctgcgaagcagctgcccagca 2322
Db 34552 GTCCAGAGCCTGAGCGCGCGGTGGCCGCCAGCAAGGGGCTGCGCAGCAGCTGCCAGCA 34493
QY 2323 cctcggagcagagtgactcagctgccccatccacgtgtccctgagtgccacccccagagcttcc 2382
Db 34492 CCTCGGAGGAGTACTCAGTGTGCCCTCCATCCACGTTGTCCCTGTGTCGGCCCCCACTTTC 34433
QY 2383 cccggcttaagcagctcctcgtgaccttaaagacatcctgagcagggccagccacccatg 2442
Db 34432 CCCGGCTTAAGCAGCTCTCGCTGAGCTTTAAAGACATCTCTGAGCAGGCGCAGCACCATG 34373
QY 2443 caactccttcagcaacagcagcaggaagcagtagatccgaagggcagcagcagcagggagcgg 2502
Db 34372 CAATCTCTTACGAACACAGCAGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGAGAGCG 34313
QY 2503 agggagcctcgggggtcccaacttcccaaggaacaattacttaggggcactctcgacc 2562
Db 34312 AGGAGGCTCGGGGGCTCCCACTTCTCTCCAGSACAAATTAATTAGGGGCATTCGACC 34253
QY 2563 attctgacaacgccaagggaggttgtaaggcagtgctcggtgtccatgggctgggtgtg 2622
Db 34252 ATTTCTGACAAACGCCAAGGAGTGTGTAAAGGCAGTGTGCGTGTCCATGGGCTGGGTGTG 34193
QY 2623 gagcgcttgagcatctgagtcaggggaagacagcttcgggggaattgcatgtacgccccca 2682
Db 34192 GAGCGTGTGAGCATCTGATCTCAGGGGAACAGCTTCGGGGGATTGCAATGTACGCCCCCA 34133
QY 2683 ctttggaggttccacccgctgtgctgtccactcctgtgccccattggccgaatgcaaa 2742
Db 34132 CTTTGGGAGTTCCACCGCTGTGCGTCCCACTCTTGTGCCCATTTGGCCGAATGCCAA 34073
QY 2743 gttctctgttagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2802
Db 34072 GGTGTCTGTGTAGCAGCAGCAGCGGAGGCGAGAGCAGTGAAGATATCTGTAGTATTCCTCT 34013
QY 2803 tcaaggaggtttacacaaagggttagagggcgagcagcagcagcagcagcagcagcagcagcagc 2862
Db 34012 TTCAAGGGAGGTTACACCAAGGGCTAGAGGGCGAGAGCCTAGGCTGTCTGGCAGCGCT 33953
QY 2863 gcagcagggagctccgggacacttgactgacgtcgctctacccctgtctctacagtcggga 2922
Db 33952 GCAGCAGGGAGCTCCGGGACACTTGAACCTGCCGTCTACCTGTCTCTCTACAAAGTCCGGA 33893
```

```
QY 2923 goactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2982
Db 33892 GCAGTGCAGGAGCAGCTCGGTACCAGAGTCCGAGACTACTACTCAACTTCCACTTGCTCTG 33833
QY 2983 gccggaccgcccccctccgcgcctcccccacccacgcctcgcacatcaagctgagagaac 3042
Db 33832 GCGGAGCCCGCCCTCCCGCGCTCCCATCCCAAGCTCGCATCAAGCTGGAGAAC 33773
QY 3043 ccgctgactacgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3102
Db 33772 CGGCTGACTACGGAGCCCTGGGCGCTGCGGCGGCGCAGTGCCTATATGGGACCTG 33713
QY 3103 gcgagcctcagtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3162
Db 33712 GCGAGCCTCATGGCGGGTGCACGGGGACCCGGTTCTGGGTACCTCAGCCGCCCT 33653
QY 3163 tctcatctctggcacactctcttcacagccagcagcagcagcagcagcagcagcagcagcagcag 3222
Db 33652 TCTCATCTCTGGCACACTCTCTTACAGCCGAAGAAGGCCAGTTCTATGGACCGTGTGT 33593
QY 3223 ggtggtgggggtggtggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 3282
Db 33592 GGTGgtgggggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 33536
QY 3283 ggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3342
Db 33535 GCGGCGGCGAGCGGGAGCTGTAGCCCTTACGGCTACACTCGGCCCTCAGGGGCTG 33476
QY 3343 ggcggccaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3402
Db 33475 GCGGCGCAGAAAGGAGCTTCAACCGACCTGATGTGTGTGATGATGATGATGATGATGATGATG 33416
QY 3403 agagtgcctatcccgagtcacactgtatcaaaagcgaatggccctgagatgagcagcagcagc 3462
Db 33415 AGAGTGCCTATCCAGTCCCACTTGTGTAAAGCGAAATGGGCCCTGTGATGATGATGATGATG 33356
QY 3463 tactcggaccttacggggacatcggtt 3490
Db 33355 TACTCCGGACCTTACGGGACATGCGGT 33328
```

RESULT 7

HUMARX LOCUS	HUMARX 3231 bp mRNA	PRI 31-OCT-1994
DEFINITION	Human androgen-receptor mRNA, complete cds.	
ACCESSION	M34233	
VERSION	M34233.1 GI:179033	
KEYWORDS	Human testis and prostate cancer cell line LNCaP, cDNA to mRNA.	
SOURCE	Human testis and prostate cancer cell line LNCaP, cDNA to mRNA.	
ORGANISM	Homo sapiens	
REFERENCE	1. (bases 1 to 3231)	
AUTHORS	Govindan,M.V.	
TITLE	Specific region in hormone binding domain is essential for hormone binding and trans-activation by human androgen receptor	
JOURNAL	Mol. Endocrinol. 4 (3), 417-427 (1990)	
MEDLINE	90258935	
FEATURES	Location/Qualifiers	
source	1..3231	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/map="Xq11.2-q12"	
gene	109..2829	
	/gene="AR"	
CDS	109..2829	
	/gene="AR"	
	/note="androgen-receptor"	
	/codon_start=1	
	/db_xref="GDB:G00-120-556"	
	/protein_id="AAA51780.1"	
	/db_xref="GI:179034"	
	/translation="MEVQLGLGRVYPRPPSKTYRGAFQNLFQSVREVIQNPGRHPPEA"	


```
QY 3616 gtcttcttcaaaagagccgtgaagggaaacacagaagtacacctgtgcgcagcagaatgat 3675
Db 1810 GTCTTCTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAGCAGAAATGAT 1869
QY 3676 tgcactattgataaattccgaaggaataattgtccattctgtcgtctcttggaataattat 3735
Db 1870 TGCACATTGATAATTCCGAAGGAATAATGTCCATCTTGTCTCTCGAAATGTTAT 1929
QY 3736 gaagcaggatgactctgtggagccgcgaagctgaagaaacttgtaactgaactacag 3795
Db 1930 GAAGCAGGAGTACTCTGGGAGCCCGGAAGCTGAAGAACTTGGTAATCTGAACACTACAG 1989
QY 3796 gaggaagagaggtcttcagcaccacagccccactcagagacacacccccagaagctgaca 3855
Db 1990 GAGGAAGGAGAGGCTTCAGACACCACACAGCCCCACTGAGGAGACAAACCCAGAGCTGACA 2049
QY 3856 gtgtcacacattgaaggctatgaatgtcagcccatcttcttgaatgtcctggagccatt 3915
Db 2050 GTGTACACATGAAGGCTATGAATGTGAGCCCATCTTCTGAATGTCTTGGAGGCCATT 2109
QY 3916 gagcaggtgagtgtgtgtgacacagacacaaacacagccccgactcctttgagccttg 3975
Db 2110 GAGCCAGGTGAGTGTGTGTGTGGACACGACAAACCAACACAGCCCGACTCCTTTGAGCGCTG 2169
QY 3976 ctctctagcctcaatgaactgggagagagacagcttgtacacgtgtggtcaagtgggccaag 4035
Db 2170 CTCCTAGCCCTCAATGAACCTGGGAGAGACAGACGCTGTACACGTGTGTCAGTGGGCCAAG 2229
QY 4036 gcttgcctggcttcgcaacttacacgtggaagaccagatggtgtcattcagctactcc 4095
Db 2230 GCCTTGGCTGGCCCTCCGCAACTTACACGTGGACGACACAGACAGTGGGTGTCATTAGTACTCC 2289
QY 4096 tggatggggtccatggtgttgcctggcgtggcgtcccttcacacgttcacacgttcacacccagg 4155
Db 2290 TGGATGGGCTCATGTGTGTGTCATGGCTGGCGATCCTTACCAATGTCAACTCCAGG 2349
QY 4156 atgtctactctgcctcctgatctgtgttttcaatgagtcacgcgtgcacaaagtcctggatg 4215
Db 2350 ATGCTCTACTTGGCCCTGTGCTGTGTTTCAATGAGTACCGCATGTCACAAAGTCCCGGATG 2409
QY 4216 tacagccagtggtccgaatgagcaccctctcagaggtttgagtggtcccaaatcaacc 4275
Db 2410 TACAGCCAGTGTGTCGAATGAGCGACTCTCAAGAGTGTGGATGGCTTCCAAATCACCC 2469
QY 4276 cccaggaattcctgtgcataagcactgctactcttcagcattatccagtggtgag 4335
Db 2470 CCCCAGGAATTCCTGTGCATGAAGCCATGCTACTCTTACGACATTATTCAGTGGATGGG 2529
QY 4336 ctgaaaaatacaaaaattccttgatgaacttcgaatgaactacatacagaactcgtatcgt 4395
Db 2530 CTGAAAAATCAAAAATTTCTTTGATGACTTCGAATGAACCTACATCAAGGAACTCGATCGT 2589
QY 4396 atcattgatgcaaaagaaaaatccacatcctgctcaagacgttctaccagctcacc 4455
Db 2590 ATCATTCGATGCAAAAGAAAAATCCACATCCTGCTCAAGAGCGCTTCACAGCTCACCC 2649
QY 4456 aagctcctggactccgtgcagcattatgagagagctgcatacagttcacttttgacctg 4515
Db 2650 AAGCTCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCATTCTTGACCTG 2709
QY 4516 ctaatcaagtcacacatggtgaagcgtggacttccggaataatgatggcagagatcatctct 4575
Db 2710 CTAAATCAAGTCACATGTTGAGCGGTGGACTTTCCGGAAATGATGCGACAGATCATCTCT 2769
QY 4576 gtgcaagtgcccaagatcctttctgggaagtcaaaccccatctatttccacaccagtgga 4635
Db 2770 GTCAAGTGCCCAAGATCCTTTCTGGAAAGTCAACCCCATCTATTTCACACCCAGTGA 2829
QY 4636 agcattggaacccctatttccccaccaccagctcgtatgccccctttcagatgtcttctgct 4695
Db 2830 AGCATTGGAAACCCCTATTTCCTCCACCCAGCTCATGCCCTTTTCAGATGTCTTCTGTGCT 2889
```

```
QY 4696 gttataactctgcactactcctctcagtcagtcgcttgggggaatttctcttattgtatgaacag 4755
Db 2890 GTTATAACTCTGCACCTACCTCTCTGCAGTGCCTTGCAGTGCCTTGGGAAATTCCTCTATTGATGACAG 2949
QY 4756 tctgtcatgaacatgttctcgaattctatttgcgtggccttttttttctctctctctctct 4815
Db 2950 TCUGTCATGAACATGTCTCTGAATCTATCTGCTGGGCTTTTTCCTCTCTCTCTCTCTCTCT 3009
QY 4816 ttcttttcttctcctcctcctatctaaacctcccatggcacccttcagacatttgcctccc 4875
Db 3010 TTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3069
QY 4876 attgtgctcctctatctgttttgatgagtggtgtatgcctttaaactctgtatgatact 4935
Db 3070 ATTGTGGCTCCTATCTGTGTTTGAATGCTGTGTATGCC-TTAAATCTGTGATGATCCT 3128
QY 4936 catatggccagtgctcaagtgtgtgttttacacagcactctctgtccagccacacaaa 4995
Db 3129 CATATGGCCAGTGTCAAGTGTGTGTTTACAC-CGATCTCTGTGCCAGCCACACAAA 3187
QY 4996 cgtttacttatttattgcaacgggaagtt 5024
Db 3188 CGGTTTACTTTACTTACCGCAAGGGAACCT 3216

RESULT 8
HUMANRE 3210 bp mRNA PRI 31-OCT-1994
LOCUS Human androgen receptor mutant gene, mRNA, complete cds.
DEFINITION M73069.1 GI:178655
ACCESSION M73069.1
VERSION 1
KEYWORDS androgen receptor.
SOURCE Homo sapiens Testis cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3210)
AUTHORS Govindan,M.V.
TITLE Specific region in hormone binding domain is essential for hormone
binding and trans-activation by human androgen receptor
JOURNAL Mol. Endocrinol. 4 (3), 417-427 (1990)
MEDLINE 90258935
FEATURES
Location/Qualifiers
1..3210
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="Testis"
/map="Xq11.2-q12"
616..2808
/gene="AR"
616..2808
/gene="AR"
/feature="mutant gene"
/codon_start=1
/db_xref="GDB:G00-120-556"
/product="androgen receptor"
/db_xref="AA51735.1"
/db_xref="GI:178656"
/translation="MQLQQQQEAVSESSSSRRAREASGAPTSKDNLTGGTSTISD
NAKELKAVSMGLGVEALEHUSPGEURGDGMVAPLIGVPPVAPVPCAPLAECKG
SLDDSGKSTEDTAESYSPKGYTKLESGSCGSAAGSSGTLFPLSLYSKS
GALDEAAVQSRDYVNFPLAGAPPPPPHPIHAKIENLIDYGSAAWAAACOCRY
GDIALAHGAGAAGPGSGPSAAASSSWHTLFTAEEQOLYPCGGGGGGGGGGGGG
EGAVAPYGYTRPPQGLAGSDFTAPDVWYPGWVSRVPYSPCTCK SEMGPMDSY
SGVGMRLFTARDHVLPIIDYFPPOKTCILGDEASGCHYCALTCGSKVFKRAAE
CKOKYLCASRNDCTIBFRKNCPSCLRKCYEAGMTGLARKLKLKLEQEGAS
STTSPEETTKLTVSHIEGECQIFLNVLEAIEPVVYVAGHNDNQPSFAALLSSL
NELGELQVHVWKAALGLRLNHYDDQMAVIQYSWGLMVFMGWRSTFTVNSRML
YFAPDLVNEYRMHKSMTSQCYMRHLQOEFGWLOITPOEFLCMKAMLIFFPELLLF
SLIPDGLKNKPFDELRLMYIKELDRIIACKRKNPTSCSRFPYOLTKLLDSVQPLAR
ELHQFTFDLLIKSHWVSVDPEPMWAEIISVQVPKILSGKVKPIYFHTQ"
BASE COUNT 703 a 930 c 875 g 702 t
ORIGIN
```

QY	2836	gagagccctaggctgctctggcagcgcttcacgacgagagctccgggacacttgaactgcgcg	2895
Db	1012	GAGAGCCTTAGGCTGTCTTGGCAGCGCTGCACAGGGAGCTCCGGGACACTTGNACTGCCG	1071
QY	2896	tctaccctgtctctctctcaagtccggagacactggacgagcagctcgtaccagatgcg	2955
Db	1072	TCCACCCTGTCTCTCAAGTCCGGAGCACTGGACGAGGACGTCGTTACCAAGATCGC	1131
QY	2956	gactatacaaatcttcacattggcttgcccggaacgcgcgccctcgcgcgctcccact	3015
Db	1132	GACTACTACAACCTTTCACATTGGCTTCTGGCCGAGCACGCGCCCCCTCGCGCGCTCCCACAT	1191
QY	3016	ccccagctcgcatacgaagtggagaacccgctggactcacgacgagcctcgtggcgctgcg	3075
Db	1192	CCCCAGCTCGCATCAAGCTGGAGAACCCTGCTGGACTACGGCAGCGCCTGGCGGCTGCG	1251
QY	3076	gcggcgadtgccgcttatgggacctggcgagcctgcattggcggggtgcagcgggaccc	3135
Db	1252	CGCGCGCAGTGC CGCTATGGGACCTTGGCGAGCCTTGCATGGCGGGGTGCAGCGGACCC	1311
QY	3136	ggttcgggtcacccctcaagcgcgcgttcctcatctcctggcacactctcttcacagccgaa	3195
Db	1312	GGTTCCTGGGTACCCCTACGCGCGCTTCCTCATCTCTGGCACACTCTCTTCACAGCGCAA	1371
QY	3196	gaaggccaagtgtgatgaaccttgttggtgggtgggtgggtgcggcgcgcgcgcgcc	3255
Db	1372	GAAGGCCAGTTGTATGGACCGTGGTGGTGGGGGGTGGT-----	1413
QY	3256	ggcggctac	3315
Db	1414	-----GCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAGCTGTAGCCCCCTAC	1467
QY	3316	ggctaacactcggccccctcaggggctgcggggccaggaagcagacttcacgcagacctgat	3375
Db	1468	GGCTTACACTCGCGCCCCCTCAGGGGCTTGGCGGGCCAGGAAAGCACTTCACCGCACCTGAT	1527
QY	3376	gtgtggttacccctggcgcatgtgtgagcadagtgcccctatcccagtcaccaacttgtcaaa	3435
Db	1528	GTGTGGTACCCCTGGCGCATGTGTACGAGTGCCTATCCCACTGCCACCTGTGTCAAA	1587
QY	3436	agcgaatggggcccttgatggtatagctactccggaaccttacgggggacatgcgtttgag	3495
Db	1588	AGCGAAATGGCGCCCTCGATTGGATAGCTACTCCGACCTTAAGGGGACATCGTTTTGGAG	1647
QY	3496	actccaggaacatatgtttgccattgactattactitccacccccaagaagacctgcctg	3555
Db	1648	ACTGCCAGGAGCAATGTTTTGCCCATTGACTATTACTTTCACCCCAAGAACCTGCTGTG	1707
QY	3556	atctgtggagatgaagcttctgggttgcactatgagctctcatactgtggaagctgcag	3615
Db	1708	ATCTGTGGAGATGAAGCTTCTGGGTGTCTACTATGAGCTCTCACATGTGGAAGCTGCAAG	1767
QY	3616	gtctottcaaaaagagcgcgtgaagggaacaagaagtaacctgtgcgcagcagaaatgat	3675
Db	1768	GTCTCTTCCAAGAGCGCGCTGAAGGGAACAGAGTACCTGTGCGCCAGCAGAAATGAT	1827
QY	3676	tgcactatgataaattccgaagaaaaattgtccatctgtcgtcttcggaaaattgttat	3735
Db	1828	TGCATAATTGATAAATTCGGAAGAAAAATTTGCCATCTTTGTGCTTCGGGAATGTTAT	1887
QY	3736	gaagcaggatgactctggagcccggaagctgaagaaacttygtaaactgaactaacag	3795
Db	1888	GAAGCAGGGATGACTCTGGGAGCGCGGAAGCTGAAGAAACTTGGTAATCTGAAACTACAG	1947
QY	3796	gaaggaagagaggtlccagcacaccagcccccaotgaggaacaaccacaagctgaca	3855
Db	1948	GAGGAAGAGAGGCTTCAGCACACACGCCCCCACCCTGAGGAGCAACCCAGAACTGACA	2007
QY	3856	gtgtcacacattgaaggctatgaatgtcacgccatctttctgaatgcctcgtgaagccatt	3915
Db	2008	GTGTACACATTTGAAGCCTATGAATGTACGCCATCTTTCTGAATCTCTGCGAAGCCATT	2067

[illegible]

|||||
Db 814 GCTGTGCGTCCACACCTCTGTGCGCCATTTGGCCGAATGCAAAAGGTTCTCTGCTAGACGAC 873
QY 2761 aqcgaggagagacactgaagataactactagatattccctttcaaggaggttacacc 2820
Db 874 ACCGACGACAGAGACACTGAAGATACTGCTGAGTATTCCTCTTCAAGGAGGTTACACC 933
QY 2821 aaagggttagaaggcgagagccttaggtgctctgagcagctgcagcaggagagctccggg 2880
Db 934 AAAGGCTAGAAAGCGAGAGCCTAGGCTGCTTGGCAGCGCTGCAGCAGGAGCTCGCGG 993
QY 2881 acactgaactgcctgtacacctgtctctataagtcctgagagcactgagcagagcagct 2940
Db 994 ACACHTGAACCTGCCGTCTACCTGCTCTCTACAAAGTCGGAGCACTGGACGAGCAGCT 1053
QY 2941 gqgtaccagagtcgcgactaactacaactttccactggctctgcccggagccgcgcacct 3000
Db 1054 GCGTACCAGAGTCGGACTACTACAACTTCCACTTGGCTCTGGCCGGACCGCCGCCCT 1113
QY 3001 ccgcgcctcccccatacccaacgctcgcatcaagctggaagcctgggaacccgcgtgactagcgaac 3060
Db 1114 CCGCCGCTCCCATCCACACCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGC 1173
QY 3061 gactgggcgtgcggcgagtcagtcacactatgggacctggcgagcctgcagtcgagcg 3120
Db 1174 GCTGGCGGCGTGGCGGCGCGCAGTCCGCTATGGGACCTGGCCGAGCCTGCATGGCCGG 1233
QY 3121 ggtgcaggggaccgggttctgggtcacacctcagcgcgcgcttctctcatctctggcacact 3180
Db 1234 GGTGACGGGACCCGGTTCTGGTCAACCTCAGCCGCGCTTCTCATCTTGGCACACT 1293
QY 3181 ctcttcacagccgaagaaagccagttgtatgacccgtgtggtgggtgggtgagtgac 3240
Db 1294 CTCTTCACAGCCGAAGAAGCCAGTTGTATGACCGTGTGGTGGTGGGGGTGGT --- 1350
QY 3241 ggcggcg 3300
Db 1351 -----GGCGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 1389
QY 3301 gctgtagcccccatacggctacactcggccccctcaggggcgcgcgcgcgcgcgcgcgcg 3360
Db 1390 GCTGTAGCCGCCCTACGGCTACATCGCGGCCCTCTAGGGGCTGGCGGGCCAGGAAAGCGAC 1449
QY 3361 ttcaacgcactgattgttgatccctgcgcactgtagcagcagagtcgccttatccag 3420
Db 1450 TTCACGCACTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGATGGCCCTATCCAGT 1509
QY 3421 cccacttgttcaaaagcgaaatgggccccctggatggatagctactcccgaccttacggg 3480
Db 1510 CCCACTTGTCTAAAAGCGAAATGGGCCCTGGATGGATGCTACTCCGGACCTTACGGG 1569
QY 3481 gacatgcgtttgagagactgcagggaccatggttttcccattgaactattactttccaccc 3540
Db 1570 GACATGCGTTTGGAGACTGCGCAGGACCATGTTTGGCCATTGACTATTACTTTCCACCC 1629
QY 3541 cagaagacctgcctgactgtgagatgaagcttctgggtgctcaactatgagagcttcaca 3600
Db 1630 CAGAAGACCTGCCTGATCTCTGGAGATGAAGCTTCGTGGGTGCTCCTATGAGGCTCTACA 1689
QY 3601 tgtggaagctgcgaaggcttcttctcaaaagagccgctgaagggaacacagaagtcacctgtgc 3660
Db 1690 TGTGGAAGCTGCAGGCTTCTTCAAAAGAGCGCTGAAGGGAAACAGAAAGTACCTGTGC 1749
QY 3661 gccagcgaatgattgcactattgataaattccgaaggaaaaattgtccattcttgc 3720
Db 1750 GCCAGCAAAATGATTGCATTTGATTAATTCGAGAGAAAAATTTGCCATCTTGTGCT 1809
QY 3721 ctccgaaattgtatgaagcagggataactctggagcccgagctggaagaaacttgg 3780
Db 1810 CTTCCGAAATGTTATGAAGCAGGGATGACTCTGGAGCCCGGAAGCTGAAGAACTTGGT 1869
QY 3781 aatctgaactacagggagggagggcttccagcaccaccagcccccactgaggagaca 3840
|||||

Db 1870 AATCTGAACACTACAGAGGAAGGAGAGGGCTTCCAGCACACACAGCCCCACTGAGGAGACA 1929
QY 3841 acccagaagctgacagtgtcacacattgaagctatgaatgtcagccccatcttcttgaat 3900
Db 1930 ACCCAGAAGCTGACAGTGCACACATTGAAGGCTATGAATGTACGCCCATCTTCTTGAAT 1989
QY 3901 gtccctggaagcctcattgagccaggtgagtgctgtgctggacagcaacaaccagcccgac 3960
Db 1990 GTCTGGAGCCCAATTGAGCCAGGTGTAGTGTGTCTGGACACGACAAACACAGCCCGAC 2049
QY 3961 tcctttgacgcttgcctctctagcctcaatgaactgggagagacagcttgtacacgtg 4020
Db 2050 TCCTTTGCACCCCTTGCTCTAGCCTCAATGAACCTGGGAGAGACAGCTGTGTACACGTG 2109
QY 4021 gtaagtggccaaagccttgctgcttcgcaacttcgcaacttacacgtggagcaccagatggct 4080
Db 2110 GTCAAGTGGGCCAAGGCCCTTGCTGGCCCTCCGCAACTTACACGTGGACGACAGATGGCT 2169
QY 4081 gtcattcagtaectcctgagtgaggctcattgtgttggccatggcgtggcgatccttcacc 4140
Db 2170 GTCAATCAGTACTCTGGATGGGCTCATGCTGTTCGCTGGCTGGCGATCCTTCACC 2229
QY 4141 aatgtcaactccaggatgctctactcttcgccccctgactgtgttttcaatgaagtcaccgtg 4200
Db 2230 AATGTCAACTCCAGGATGCTTACTTCGCCCTTGATCTGTGTTTCAATGAGTACCGCATG 2289
QY 4201 cacaagtcgcggatgtacagccagtggtccgaatgaagcaccctctcctcaaggtttgga 4260
Db 2290 CACAAGTCCCGGATGTACAGCCAGTGTCCGAATGAGGACCTCTCTCAAGAGTTTGA 2349
QY 4261 tggctccaaatcacccccccaggaaattcctgtgcatgaaagcactgactcttcacgatt 4320
Db 2350 TGGCTCCAATCACCCCCAGGAATTCCTGTGCATGAAAGCATGTCTCTTCAGCAT 2409
QY 4321 attccagtgagtggtgaaataacaaaattcttctgactgaactcctcaatgaactacac 4380
Db 2410 ATTCAGTGGATGGGCTGAAATAATCAAAAATTCCTTGTGAATTCGAATGAACATACATC 2469
QY 4381 aagaaactcgatcgtatcattgcatgcaaaagaaaaatccccacatcctgctcgaagcgc 4440
Db 2470 AAGAACTCGATGTATCATTTGCATGCCAAGAAAAAATCCACATCTCTGCTCAAGACG 2529
QY 4441 ttctaccagctcacaagctcctggaactccgtgcagcctattgcgagagagctgcacag 4500
Db 2530 TTCTACCAAGCTCACAAGCTCCTTGACTCCGTGCATCCTATTGGGAGAGAGTGCATCAG 2589
QY 4501 ttcaactttgacctgctaatacaagtccacatgctgaagctggacttccggaataatg 4560
Db 2590 TTCACCTTTGACCTGCTAATCAAGTCACATGGTGAAGCTGGACTTTCCGGAAATGATG 2649
QY 4561 gcagagatcctctgtcgaagtgcccaagatccttcttggaagtcgaagcccatctat 4620
Db 2650 GCAGAGATCATCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTAT 2709
QY 4621 ttccacacccagtggaagcattggaaacccctatttccccccccagctcagtcaccccttc 4680
Db 2710 TTCCACACCCAGTGAAGCATTTGAAACCCCTATTTCCCAACCCAGCTCATGCCCTTTC 2769
QY 4681 agatgttctgctgttataactctgcactactcctctcagtcgcttgggaattt 4738
Db 2770 AGATGTCTCTGCTGTATATACTCTGCACACTCTCTGAGTGCCTTGGGGAATTT 2827

RESULT 12

MFU94179
LOCUS MFU94179 2821 bp mRNA PRI 16-SEP-1999
DEFINITION Macaca fascicularis androgen receptor mRNA, complete cds.
ACCESSION U94179
VERSION U94179.1 GI:3861480
KEYWORDS
SOURCE crab-eating macaque.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Cercopithecinae; Macaca.	
REFERENCE	1 (bases 1 to 2821)
AUTHORS	Chong,C.S., Kempainen,J.A. and Wilson,E.M.
TITLE	Evolution of the primate androgen receptor: a structural basis for disease
JOURNAL	J. Mol. Evol. 47 (3), 334-342 (1998)
MEDLINE	98404153
PUBMED	9732460
REFERENCE	2 (bases 1 to 2821)
AUTHORS	Chong,C.S., Kempainen,J.A. and Wilson,E.M.
TITLE	Direct Submission
JOURNAL	Submitted (18-MAR-1997) Laboratories for Reproductive Biology, University of North Carolina at Chapel Hill, CB 7500, MSRB, Rm 370, UNC-CH, Chapel Hill, NC 27599, USA
FEATURES	Location/Qualifiers
source	1..2821
	/organism="Macaca fascicularis"
	/db_xref="taxon:9541"
CDS	1..2688
	/codon_start=1
	/product="androgen receptor"
	/protein_id="AAC73050.1"
	/db_xref="GI:3861481"
	/translation="MEYQLGLRVYPRPSKTVRGAFQNFILFSQVREIQNPGRPHPEAASAAPCAQLQQQQQQQEQSPRQQQQQQQEDGSPOAHRGPTGYLVLDREQPSQPSQTADECHPERGCVPEGCAAVAAKGLPQQLPAPDEDDSAAPTSLSLGPTFFGLSSCSTLKDILFSASTNQLLQQQQQQLAQSSGSRAREASGAPTSSKNDLPGGTSTIDSSAKELKAVSVSMGLVQALHSPGELRGDCMYAVLGPVPAVRPTPCAPLAECGSLALDDASGKSTEDYTFSPFKGGT KGLESGSCGSAAGASGOTLEPSTLSLYKSGALDEAAQSGRDYTFNPLAAGPPPPPPPHPHARIKLENP LDYGSAAAAAQAQCRGDLASLHGAGAAGPGSPSAASSSWHTLFTAEQOILYPCGGGGGGGGGGGAGGAEAVAVAPXYTRPPQGLAGQEDFTADPMVYPGMSVRVPYSPTCVKSEMGPMWDSYSGPYGDMLETAARDHVLVDYVFPQKTLCLIGDEASGCHYGALTCGCKVFFKRAAGKQYKLCSARNDCITDKFRKNCPSRLKRYEAGMTLGARKLKLGNLKIQDEGEASTTSPTEAQLKTVSHIEGYEQDPIFNVLIEPGVVVCAGDHNNQPDFAALISLNLGLBRLVHVHWKAKALPFGFNLHWDQMAVIQYSWGLMVFAMGRSFTNVNRSRLYFAPDLVNEYRMHRSRYSCQVRNHDLSQEFGLQITPOEFLCMKALLFSIIIPVDGLKNOFFEDLMNYIKELDRITACKRNKPTCSRFRYOLTLLKDSVQPTARELHQFTDLELIKSHVSVNDFPMMAEILISVQPKILSGKVKPIYFTQ"
BASE COUNT	623 a 829 c 799 g 570 t
ORIGIN	
Query Match	51.4%; Score 2612.4; DB 9; Length 2821;
Best Local Similarity	95.3%; Pred. No. 0;
Matches 2771; Conservative	0; Mismatches 46; Indels 91; Gaps 4;
QY 1864	atggaagtcagttagcgtgggaaggtctacctcggcgcgctccagacctaaccga 1923
Db	
1	ATGAGGTGCAGTTAGGGCTGGGAGGGTGTACCTCGCGCGCGTCCAGAGCCTACCGA 60
QY 1924	ggagcttccagaatctgttccagagcgtgcggaagtgcacagaaccgcggccccag 1983
Db	61 GGAGCTTTCAGAAATCTGTTCCAGAGCGTGGCGGAAGTGATCAGAACCCGGCCCCAGG 120
QY 1984	caccagagggccgagcgagcacctcccgcgccagtttgcgtgctgcagcagcag 2043
Db	121 CACCCAGAGCGCGAGCGCAGACACCTCCCGGGCGCAGTTT----- 161
QY 2044	cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2103
Db	162 -----GCAGCAGCAGCAGCAGCAG 180
QY 2104	cagcagagagactgccccagg-----cagcagcagcagcagcaggttgaggtcttccc 2160
Db	181 CAGCAAGAAACTAGCCCCCGGCACACGACGACGACGACGAGGTGAGTGTCTTCCC 240
QY 2161	caagccatcgtagagccccacagactacctgttgcctgagtagaacaagaacctca 2220
Db	241 CANAGCCATCGTAGAGCCCCCAGAGCTACCTGGTCTGATGAGGAACAGCAGCTTCA 300
QY 2221	cagccgcagtcggccctggagtgcacccccagagaggttgcgtccagagcctgaqcc 2280

Db 1354 GCTGTAGCCCTTACGGCTACACTCGGCCACCTCAGGGGCTGGCGGCCAGGAAGCGCAC 1413
Qy 3361 ttacacgcacatgatgtgtgtacccctggcgcatgtgtgagcagatgcccctatccccagt 3420
Db 1414 TTCACCGCACCTGATGTGTGTCACCTGGCGGCATGTTGAGCAGAGTGCCCTATCCCACT 1473
Qy 3421 cccacttgttcaaaagcgaataatgggcccctggatgagatgactactccgacaccttaagg 3480
Db 1474 CCCACTTGTGTCAAAAGCGAGATGGGCCCTGGATAGTACTCTCCGACACCTTACGGG 1533
Qy 3481 aacatgcttggagactgcccagagaccatgttttggccattgactattactttccacccc 3540
Db 1534 GACATGCGGTTGGAGACTGCCAGGACCATGTTTGGCAATTGACTATTACTTTCCACCC 1593
Qy 3541 cagaagacgctgctgatctgtggagatgaagcttctgggtgtcactatggagctctcaca 3600
Db 1594 CAGAAGACTGCTGATCTGTGGAGATGAAGCTTCTGGGTGTCATATGGAGCTCTCACA 1653
Qy 3601 tgtggaagctgcaaggtctcttcaaaagaccgtgaaggagaaacagaagtacctatgc 3660
Db 1654 TGTGGAAGCTGCAAGGTCTTCTCAAAAGAGCCGCTGAAGGGAACAGAAATGACTGTGT 1713
Qy 3661 gccagcagaatgattgacattatgaataatccgaaggaataatgtccatcttctgtcgt 3720
Db 1714 GCACGAGAAATGATTGACATATTGATAATTCGAAAGGAAATTTGCCATCTTGCCT 1773
Qy 3721 cttcggaatgtttgaagcagggatgactctgggagcccggaagctgaagaaacttggt 3780
Db 1774 CTTGGAATGTTATGAAGCAGGATGACTCTGGAGCCGCGAAGCTGAAGAAACTTGGT 1833
Qy 3781 aatctgaaactacagaggaagaggcttccagcaccaccagcccaactgagagaca 3840
Db 1834 AATCTGAAACTACAGGAGGAGAGAGGCTCCAGCACCCAGCCCTGAGGAGACA 1893
Qy 3841 accagaagctgacagtgcacacatgaagctatgaatgtcagcccatctttctgaat 3900
Db 1894 GCCCAGAACTGACAGTGTACACATTGAGGCTATGAATGTACGCCCATCTTCTGAAT 1953
Qy 3901 gtccctggaagccattgagccaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3960
Db 1954 GTCTTGGAAAGCCATTGAGCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2013
Qy 3961 tcccttgagccttctctctagctccaatgaactgaactgagagagacagctgtacacatg 4020
Db 2014 TCCTTCGACGCTTGTCTCTTCTACCTTCAATGAACCTGGGAGAGACAGCTTGTACATGT 2073
Qy 4021 gtcaagtggccagccttgcctgcttccgcaacttaccagctgagacagatggct 4080
Db 2074 GTCAAGTGGCCAAAGGCTTGCCTGCTTCCGCACTTACACGTGACGACACAGATGGCT 2133
Qy 4081 gtcattoagtaactcctggatggggctcatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4140
Db 2134 GTCAATTGACGTACTTCTGATGGGCTCATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2193
Qy 4141 aatgtcaactccagatgctctacttctgcccctgactgtgtgtgtgtgtgtgtgtgtgtgtgt 4200
Db 2194 AATGTCAACTTCCAGGATGCTTCTACTTTGGCCCTGATCTGTGTTTCAATGAGTACCCCATG 2253
Qy 4201 cacaagtccccgagatgtacagccagtggtgtccgaatgagggcacctctctcaagagtttga 4260
Db 2254 CACAAGTCCCGGATGTACACCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGA 2313
Qy 4261 tggctcacaatacccccccgaggaaattctctgtgataagaaagcactgtactcttcagcatt 4320
Db 2314 TGGCTCCAAATCAACCCCCAGGAATTCCTGTGATGAAAGCGCTGTACTCTTTCAGCAT 2373
Qy 4321 attcagtgatggctggaataatcaaaaattcttggatgaacttgaatgaactacalc 4380
Db 2374 ATTCAGTGGATGGCTGGAATAATCAAAAATCTTTGATGAATTCGAATGAATACATC 2433
Qy 4381 aaggaaactcagctatcatctgcatgcaaaagaaaaataccccacatctctgctcaagacgc 4440
Db 2434 AAGGAACCTCGATCGTATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2493

Qy 4441 ttctaccagctcacaagctcctggaactccgtgagcctattgagagagctgcatcag 4500
Db 2494 TTCTACCACTCACCAGCTCCTGGACTCCGTGCGACCTATTTCGAGAGAGCTGCATCAG 2553
Qy 4501 ttcaactttgacctgtaactcaagtccacatggtgagcgtggactttccggaataatgatg 4560
Db 2554 TTCACTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCGGAATATGATG 2613
Qy 4561 gcagagatcatctctgtgcaagtgcacagatcctttctgggaaagctcaagccatctat 4620
Db 2614 GCAGAGATCATCTCTGTGCAAGTGCCCAAAATCCTTCTTGGGAAAGTCAAGCCCATCTAT 2673
Qy 4621 ttccacaccagtgagcattgaaa-cctattttcccccacccacagctcatcccccttt 4679
Db 2674 TTCCACACCCAGTGAAGCATTTGAAATCCCTATTCTCACCACCATCATGCCCTTTT 2733
Qy 4680 cagatgtcttctgctgttataaactctgcactactcctctgcagtgcccttggggaatttc 4739
Db 2734 CAGATGCTTCTGCCCTGTATAACTCTGCACACTACTCTCTGCAGTGCCTTGGGGAATTC 2793
Qy 4740 ctctattgatgtcacagctctgcatgaac 4767
Db 2794 CTCTATTGATGTACAGTCTGTCTCATGAAC 2821

RESULT 13
PHU941176
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
baboon.
Papio hamadryas
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
1 (bases 1 to 2769)
Choong,C.S.; Kempainen,J.A. and Wilson,E.M.
Evolution of the primate androgen receptor: a structural basis for
disease
J. Mol. Evol. 47 (3), 334-342 (1998)
98404153
732460
2 (bases 1 to 2769)
Choong,C.S.; Kempainen,J.A. and Wilson,E.M.
Direct Submission
Submitted (18-MAR-1997) Laboratories for Reproductive Biology,
University of North Carolina at Chapel Hill, CB 7500, MSRB, Rm 370,
UNC-CH, Chapel Hill, NC 27599, USA
Location/Qualifiers
1. .2769
/organism="Papio hamadryas"
/db_xref="taxon:9557"
1. .2688
/codon_start=1
/product="androgen receptor"
/protein_id="AAC73047.1"
/db_xref="GI:3861473"
/translation="MEVQLGLGRVYPRPPTKTYRGAFQNLFSQVREVIQNPGRHPDEA
ASAAPCHPCVPEPCAAVAAGKLPOLAPPDEDDSAAPSTLSLGLPTFGLSSC
SADLKDLSEASTMOLLQOOQAVSEGSRRARASGAPTSKKNYLGSTIIDS
AKELCAVSVSMGLGVLEHLSFGQLRQDCMYAPVIGVPVAPRPTCPALAEKGS
LLDDSAKSTEDTAEYSPFKGTYKLGLESLGCSGAASGSGTLELSTLSLYKSG
ALDAAAYQSRDYNFPLALAGPPPPPPHARIKLENPLDYGSAWAAAAQCRIG
ELASLHGAGAAGSGSPSAASWHTLEAEGLYPCGGGGGGGGGGGAGGEEA
GAVAPGYTRPQGLAGDEFTAPDVMYGMVSRVPSPCTVKSEMPWMDYSYG
PYGDMRLTARDHVLPTDYFPQKTCICGDEASGCHYALTCGCKVFFKRAEKG
QKYLCAARNOCITDKFRKNCPCRLKCYEAGMTLGARKLKLGKLNKLOEEGASST
TSPTTEAOKLTVSHIEYCQIFLNVLEAIEPGVVCAGHDNNQDPSPFALLSSLINE
LGERQLVHVVKWAKALPFRNLHVDDQMAVLIQYSWMGLVMFAMGWSFTNVNRSRLIF

APDLVFNEYRHHKSRMYSCQVRNRHLSQEFGLWQITPQEFCLMKALLFSIIPVDGLK
NQKQFDELRRNYIKELDRITACKRNKPTSCSRFFYQITKLDSVQPIARELHQFTFDL
LIKSHWYSDVFPENMAEIIISVQPKILSGVKPIYFHTQ*

BASE COUNT 612 a 815 c 788 q 554 t

Query Match 50.8%; Score 2579.4; DB 9; Length 2769;
Best Local Similarity 95.4%; Pred. NO. 0;
Matches 2722; Conservative 0; Mismatches 46; Indels 85;

QY	1864	atgaaatgcaattagggctgggaagggtctacccctcggcgcgcgtccagadactaccga	1923
Db	1	ATGGAAGTGCAGTTAGGGCTGGGAGGGTTACCTTCGGCGCGCTCAAGACCTACCGA	60
QY	1924	ggagctttccagaatctgttccagagcgtgcgcgaagtgtaccagaaccgcggcccccag	1983
Db	61	GGAGCTTTCCAGAATCTGTTCCAGAGCGTGGCGAAGTGATCCAGAACCCGGGCCCCAGG	120
QY	1984	caccagagccgcgcagcgcagacaacctccggcgcagtttgcgtgcgcagcagcag	2043
Db	121	CACCACAGGCGCGAGCGAGACACCTTCGGGGGCCAGTTT-----	161
QY	2044	cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag	2103
Db	162	-----GCAGCAGCAGCAGCAGCAG	183
QY	2104	cagcaagagactgcccccagcagcagcagcagcagcagcagcagcagcagcagcagcag	2163
Db	184	CAGCAAGAAACTTAGCCCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	243
QY	2164	gcccatcgtagagcccccagagctacctggttcctggatgagaaacagcaaccttcacag	2223
Db	244	CCCCATGCTAGAGGCCCCACAGGCTACCTGGTCTGGATGAGGAACAGCAGCCTTCACAG	303
QY	2224	ccgcagtcggccctggagtgccaccocccagagaggttgcgtccagagccttcgagcgcgc	2283
Db	304	CCTCAGTCAGCCCCGGAGTGCCACCCCGAGAGAGTTGGCTGCCAGAGCCTTGAGCGGCC	363
QY	2284	gtggccgcagcaagggtctgcgcagcagcagctccagcagcctccgacagagatgactca	2343
Db	364	GTGGCCCGCGCAAGGGCTGCCGCGACAGCTGCCAGACCTCCGACACAGGATGACTCA	423
QY	2344	gctgccccatccacgttgtccctgtctgggcccactttcccgcgtttaaagcagctgtcc	2403
Db	424	GCTGCCCATCCACGTTGTTCTGCTGGCCCCACATTTCCCCGGCTTAAGCAGCTGCTCC	483
QY	2404	gctgaacctaaagacatctctgaagcagcagcagcagcagcagcagcagcagcagcag	2463
Db	484	CGCGACCTTTAAGACATCTGTGACGAGGCGACGACCATCACTCTTCAGCAACAGCAG	543
QY	2464	caggaagcagtatccgaagcagcagcagcagcagcagcagcagcagcagcagcagcagc	2523
Db	544	CAGGAAGCAGTATCCGAAGCAGCAGCAGCGGGAGACGAGGGGCTCGGGGGCTCCC	603
QY	2524	acttcctcaaggaacaattacttaggggcacttcgaccatttctgacaacgcccaaggag	2583
Db	604	ACTTCTCCAAGGACAAATTACTTAGGGGCACTTCGACCATTTCTCAGACGCCCAAGGAG	663
QY	2584	ttgtgtaagcagcagtgctcgtgtccatggcctgggtgtgagggcgttgagacatctgagt	2643
Db	664	CTGTGTAAAGGCAGTGTGGGTGTCCATATGGGCTTGGGTGTGGAGCGCTTGAGCATCTGAGT	723
QY	2644	ccaggggaacagcttcggggggatgcatgtacgcccccaacttttggaggttccaccgct	2703
Db	724	CCAGGGGAACAGCTCGAGGGGATTTGCATGTACGCCCCACATTTTGGGAGTTTCCACCCCGCT	783
QY	2704	gtgcgtcccaactcctgtgccccatggccgaatgcaaggttctctgctagacagcagc	2763
Db	784	GTGCGTCCACATCCCGTGTGCCCATTTGGCCGAATGCAAGAGTTCTCTGCTAGACACAGC	843
QY	2764	gcaggcaagagcactgaagatactgtagtattcccctttcaagggaaggttacacaaa	2823

QY	3904	ctggaagcattgagccagggtgtagtgtgctggacacgacaacacagcccgactcc	3963
Db	1957	CTGGAGCATTGAGCCAGGTGTGGTGTGCTGCATGACAAACACAGCCGACTCC	2016
QY	3964	tttgcagccttgctctctagcctcaatgaacatgggagagagacagctgtgtacacgttggtc	4023
Db	2017	TTTCGAGCCTTCTCTCTAGCCTCAATGAACATGGGAGAGACAGCTGTGTACATGTGGTC	2076
QY	4024	aagtgggcaaggccttgccggcttcgcgaacttacacatgaacgaccagatggctgtc	4083
Db	2077	AAGTGGGCCAAGGCCCTTCCTGGCTTCGCAACCTTACACGTGGACGACCATGGGTGTC	2136
QY	4084	attcagtaactcctggatggggcctcatggtgtttgccatgggtggcgatccttccaccaat	4143
Db	2137	ATTCACTACTCTCGATGGGCTCATGGTGTTCGCCATGGCTGGCGATCCTTCCACCAAT	2196
QY	4144	gtcaactccagatgctcttaactctgcccctgatctgtgttttcaatgagtaccgcatagcac	4203
Db	2197	GTCAACTCCAGGATGCTCTACTTTTCCCTCATCTGGTGTTCATGTAGTACCGCATCCAC	2256
QY	4204	aagtccggatatacagccagtggttcgaatgagggcacctctctcaagagtttgatgg	4263
Db	2257	AAGTCCGGATGTACAGCCAGTGTCTCGAATGAGGCACCTTCTCAAGAGTGTGGATGG	2316
QY	4264	ctccaaatcacccccccaggaaattcctgtgcataaggaacactgctactcttcacgattatt	4323
Db	2317	CTCCAANTCACCCCCAGGAATCTCTGTGCATGAAGCGTGTCTTCTTCACGATTAAT	2376
QY	4324	ccagtggaatgggctgaaaaatcaaaaattctttgatgaacttcgaatgaactacatcaag	4383
Db	2377	CCAGTGGATGGGCTGAAAAATCAAAAATTTCTTGATGAATTCGAATGAACATACATCAAG	2436
QY	4384	gaactcgatgctatgcatgcatgcaaaaagaaaaatcccaatcctcctcctcaagaacgttc	4443
Db	2437	GAACTCGATCGTATCATTCGATTCGCAAAAAGAAAAATCCACATCCTCTCTCAAGCGTTC	2496
QY	4444	taccagctcaccaagctcctggactccgtgcagcctattgcgagagagctgcatacagttc	4503
Db	2497	TACAGCTCACACAGCTCTCTGGACTCCGTGCAGCCTATTGGAGAGAGCTGCATCAGTTC	2556
QY	4504	actttgacctgctaatcaagtacacatgctgagcgtggactttccggaaatgatggca	4563
Db	2557	ACTTTTGACCTGCTAATCAAGTCACACATGCTGAGCGTGGACTTTCCGGAATGATGGCA	2616
QY	4564	gagatcatctgtgcaadgcccagaatcctttctgggaagaagcaagcccatctatttc	4623
Db	2617	GAGATCATCTGTGTGAAGTGCCCAAGATCCCTTTCTGGGAAGTCAAGCCCATCTATTTC	2676
QY	4624	cacaccagtgagcatttgaaa-ccctatttccccccccccagctcatgccccctttcag	4682
Db	2677	CACACCCAGTGAGCATTTGGAATCCCTATTTCCTACCCCGAGCTCATGCCCTTTTCAG	2736
QY	4683	atgtcttgcctgttataactctgcaacttc 4715	
Db	2737	ATGTCTTCTGCCTGTATAACTCTGCACTACCC 2769	
RESULT	14		
109511			
LOCUS	109511	4135 bp	PAT 02-DEC-1994
DEFINITION		Sequence 3 from Patent WO 8909791.	
ACCESSION	109511		
VERSION	109511.1	GI:587783	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE			
AUTHORS	1 (bases 1 to 4135)		
TITLE	French, F.S., Wilson, E.M., Joseph, D.R. and Lubahn, D.B.		
JOURNAL	DNA ENCODING ANDROGEN RECEPTOR PROTEIN		
FEATURES	Patent: WO 8909791-A 3 19-OCT-1989;		
	Location/Qualifiers		

```

source      1. .4135
            /organism="unknown"
BASE COUNT  979 a 1147 c 1082 g      924 t      3 others
ORIGIN

Query Match      48.8%; Score 2478.6; DB 6; Length 4135;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 3407; Conservative 0; Mismatches 665; Indels 306; Gaps

Qy 758 atcccgaggagccagcttgcggagagcgggaacggtccggagcagcccgagagcgaga 817
Db 1 ATCCCTAGGAGCCAGCCTGCTGGAGAACCCAG-AGGGTCGGAGCAAAACCTGGAGGCTGA 59

Qy 818 ggagcgacagagggaaaaagggcccnagctagccgctccagtcgctgtgtacagnagccgaa 877
Db 60 GAGGSCATCAGAGGGGAAAAGA--CTGAGTTAGCCACTCCAGTGCATACAGAAGCTTAA 117

Qy 878 -ggagcaccacgccagcccccagccgggtccagcgacagcnaacgcctcttcgancggt 936
Db 118 GGGACATACCACGCAGCCGCCAGCCACAGCCAAACGCCTGTTGCAGAGCGCGGCT 177

Qy 937 tcgaagccgcgcgcc-ggagctgcccttctcttcggtgaagtgtttaaagctctaa 995
Db 178 TCGAAGCGCGCGCCCAAGAAGTGCCCCCTTCCTCTTCGGTGAAGTTCTTAAAGCTCGGG 237

Qy 996 agactcggaggaagcaagaaaagtgcctggttagaactgacggtgccttgcctctccc 105
Db 238 ACACTCGGAGAACGCAAGAAAGTTCGGTAGGACT-ACGACTGCCTTTGTCTCCTCC 296

Qy 1056 tctcaccccgcctccccccaccctgccttccccctcccccgcttcttctccccgagc 111
Db 297 CTCCTATCCC-----CTACCCCTCCTGGGTCCCTCTCCTCTGAGC 335

Qy 1116 tgcctagtcggtactctctcagcaacccccctcacaacccctctccccacccgcgcc 117
Db 336 GGACTAGCAGCGATTC-----CTGGCCAGCC 361

Qy 1176 cgcgccgctgcgccagcgnatcagcgcgagttgagagagtaacctcccttggctg 123
Db 362 CTCCTCCCTACACCACAGACTCTGCCAGCCAGTTTGACAGAGGTAACTCCCTTTGGCTG 421

Qy 1236 cgagcgggcagagctagctgcacattgcgaagaaggctcttagag-caggcgactgggg 129
Db 422 AAAGCAGACGAG-CTTGTTGCCATTGGAGGGAGGCTTTTGGAGGCCAGAGACTGAGG 480

Qy 1295 agcggcttcagcaactgcagccagcaacnccgctggttagctgcacgcgagagaaccc 135
Db 481 AGCAACAGCAGCGCTGGAGAGTCCCTGATTCAGAGTT----- 516

Qy 1355 tgtttccccacactctctccacctcctcctcctcc-----caccccgagtcgagcc 141
Db 517 -----CTCCCCCTGCACCTCCTACTGCGCGCCCTCACCCGTGTGTGCAGCT 565

Qy 1413 agagatcaaaagatgaaggcagtcaggttcttcagtagcccaaaaaaacaacaaacaa 147
Db 566 AGAATTGAAAGATGAAAAGACAGTTGGGCTTCAGTAGTCGAAAGCAAAAC----- 617

Qy 1473 acaaaaaagccgaaaataaaagaaaaagataataaactcagttcttatttcacactactc 153
Db 618 -----AAAAGCAAAAGAAAAACAAAAGAAATAGCCAGTGTCTATTATGACCTGTCTC 672

Qy 1533 agtgacactgaatttgaaggtgagagatttgtttttctttttaaagatctggggcatc 159
Db 673 AGTGACATTGACTTTGGAAGGCAGAGAAATTTCTTCCCCCCTCAGCTCAAGCTTTGAGCAT 733

Qy 1593 ttgtgaatctacccttcaagtattaaagacagactgtgagccttagcagggcgagatttg 165
Db 733 CTTTAAATCTGTTCTCAAGTATTAGGGCAAAACTGTGAAACTAGCAGGGCAGATCCTG 799

Qy 1653 tccacgctgtgtcttctctcgcagcagactttgaggtgtgcagagcgtttttgtcgtggt 171
Db 793 TCTACGGGTGCTCTCCCTTTTACAGAGACTTTGAGGTATCTGGCGCTCTCCCCCTC- 851

```

QY	1713	tgctccgcgaagtttctctctctgtggagcttcccgcaaggtgggcagctagctgctgcagcagct	1772
Db	852	---CCGTCAAAGTTTCTCCCTG6GAGCTTCGCGCAGGTGGCAGCTAGCTGCGATACT	907
QY	1773	accgcatacatcagcctgttgactctctctgagcaagagaagggagcgagctaaagg	1832
Db	908	ACATCAT-----CAGTCAGTAGAACTCTTTCAGAGCAAGAGACAGAGGCGCAGGATAAGGG	962
QY	1833	aagtagtggaagattcagccaagctcaaggatggaagtgcagcttaggctgggaaggggt	1892
Db	963	AATTCCGTGGAAGCTAGAGACAAGCTAAAGGATGGAGGTGCGATTAGGCTGGGAAGGGT	1022
QY	1893	ctacccctggccgcgttccaagacctacagagagcttccagagagcttccagaatatctgttccagagcgt	1952
Db	1023	CTACCCACGGCCCCGCTCCCAAGACCTATCTGAGAGAGCGTTTCCAGAATCTGTTCAGAGCGGT	1082
QY	1953	gcgcgaagtgatccagaacccgggcccagggcacccagagggccgcgagcgagcacctcc	2012
Db	1083	CGCGAAGCGGATCCAGAACCCGGCCCCAGGACACCTTGAGGCCGTAGCATAGCACCTCC	1142
QY	2013	cggcgcagttgtcgtcgtcgcagcagcagcagcagcagcagcagcagcagcagcagca	2072
Db	1143	CGGTGCTGTTTACAGCAGCGCGAGGAGACTAGCCCCCGGC-----	1183
QY	2073	gcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagca	2132
Db	1184	-----GGCGCGCGCG	1193
QY	2133	gcagcagcaggtgagatggtttcttcccacagcccatcgttagagggcccacagctacct	2192
Db	1194	GCACGACCCCTGAGGATGGCTCTCTCTCAANNNCACATCAGAGCCACACAGCGCTACT	1253
QY	2193	ggtcctggatgagaaacagcaaccttcacagccgcagtcggccctggagtgcaccccga	2252
Db	1254	GGCCCTCGAGGAGGAACAGCAGCGCTTTCAGCAGCAGTCAGCCTCCGAGGGGCCACCCCTGA	1313
QY	2253	gagaggttgctccagagaccttgagccgcgcgtggcccccagaggggtgcgcagca	2312
Db	1314	GAGCGGTGCTCTCCCGGAGCGCTGGAGCTGCCACGCTCTCTGGCAAGGGGTGCTCGCGAGCA	1373
QY	2313	gctgcagcaccctccgcgcagagatgactcagctgccccatccacgttgttccctgtcggg	2372
Db	1374	GCCACAGCTCTCCAGATCAGGATGACTAGCTGCCCCATCCACAGTTGTCCCTTACTTGGG	1433
QY	2373	cccaacttccccggcttaagcagctgctccgcgtgaccttaagacatccctgagcgaggc	2432
Db	1434	CCCCACTTCCCAGGCTTAAGCAGCTGTCTCCGCAGACATTAACACATCTCTGACGAGGC	1493
QY	2433	cagcaccatgcaactcctt-----	2451
Db	1494	CGGCACCATGCAACTTCTTTCAGCAGCAGCAACAGCAACAGCAGCAGCAGCAGCAGCA	1553
QY	2452	-----cagcaacagcagcaggaagcagatccgaagcagcagcagcagcgagagc	2501
Db	1554	GCAGCAGCAGCAGCACACACAGAGGTTATATCCGAAGGCGAGCAGCAGCTGTAGAGC	1613
QY	2502	gagggagcctcgggggtccccacttctccaaagacaattacttagggggcacttcgac	2561
Db	1614	AAGGAGGCCACTGGGGCTCCCTCTTCTCCTCAAGATAGTTACCTAGGGGGCAATTTCGAC	1673
QY	2562	caattctgacaacgccaaggaattgtgttaagcgagtgctcgtgtccatggggctgggtgt	2621
Db	1674	CATATCTGACATGTCACAGGAGTTGTGTAAAGCAGTGTCTGTGTGCCATGGGTGGGTGT	1733
QY	2622	ggagcggttggagcacttgactcaggggaaacagcttcgggggattgcatgtacgcccc	2681
Db	1734	GGAGCAGCTGGAAACATCTGAGTCCAGGGGAGCAGCTTCGGGGGCGACTGCTATTCGCTC	1793
QY	2682	acttttgggagttccaccgcgtgtcgtcccaactccttgtgcccaattggccgaatgcaa	2741
Db	1794	GCTCCTGGGAGGTCCACCGCGGTGCGTCCCACTCCTTGTGGCCCTCTGGCCGAATGCAG	1853

QY	2742	aggttctctgtagacgacgagcgagcgagcactgaagtaactgctgagttatcccc	2801
DB	1854	AGGTTTCCCTGGACGAAGGCCCGGGCAAAAGCAGCTGAAGAGACTGCTGAGTATTCCTC	1913
QY	2802	tttcaaggagaggtttacacccaagggtctagaaggcgaagacctaggctgctctgagcgcg	2861
DB	1914	TTTCAAGGGAGGGTTACGCCAAAGGGTTTGAAGGTGAGACTCTGGCGTCTCTCGGCGAGCAG	1973
QY	2862	tgacgagggagctcccggggacacctgaactgccgtctaccctgtctctctctcaaatgcgg	2921
DB	1974	TGAACACAGTAGCTCTGGGACACTTGAGATCCCGTCTCTCACTGTCTGTATTAAGTCFNG	2033
QY	2922	agcactggacgagcgactgctgctaccagagtcgcygactactcaaatcttccactggctct	2981
DB	2034	AGCAGTAGCAGGCGACGACATACAGAATCGGACTACTACAATTTCCGCTCGCTCT	2093
QY	2982	ggcgagccgcccgcctccgcgcctccccatccccagctcgcatcaagcttggagaa	3041
DB	2094	GTCCGGGCGCGCACCCCGCGCCCTACCCATCCACACGCCCGCATCAAGCTTGGAGAA	2153
QY	3042	ccgcgtgactacgagcgacctggcgctgcgcgcgcgcagtcgcgctatggggacct	3101
DB	2154	CCGCTCGGACTACGCGACGCGCTGGGCTGGGCGCGCAGCGCAATCGCGCTATGGGGACTT	2213
QY	3102	ggcagccctgcatggcggggtgacgggagcccggttctgggtcacccctcagcgcgcg	3161
DB	2214	GGCTAGCCTACATGGAGGGAGTGTAGCGGAGCCACGACACTGGATGCCGCCACGACCGC	2273
QY	3162	ttctcatctctggcacactcttccacgcggaagggccaagttgtatgacccgtgtgg	3221
DB	2274	CTCTTCTTCCCTGGGCATACTCTCTTCACAGCTGAAAGAGGCCAATTTATATGGGCC	2327
QY	3222	tggttggtgggggtggtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg	3281
DB	2328	-----AGGAGCGGGGGCGGCACGACG	2348
QY	3282	cggcgcgcggaagcgggagctgtagccccctacgctacactcgccccctcaggggct	3341
DB	2349	TAGCCCAAGCGATGCTGGGCGCTGTAGCCCGCTATGGCTACACTCGGCCCGCTCAGGGGCT	2408
QY	3342	ggcgggccagaaagagacttcaacgcacactgaatgtgtgtggtacctggcggaatggtgag	3401
DB	2409	GGCAAGCGCAGAGGGTACTTCTCTGCCCTCTGAAGTGTGGTATCCTGTGGAGTGTGAA	2468
QY	3402	cagagtgcctatccacgtcccaacttgtctaaagcgaaatggccccctggatggatag	3461
DB	2469	CAGAGTCCCCATCCGAGTCCCGAGTGTGTTTAAAAGTGAATGGAGCTTGGATGGAGAA	2528
QY	3462	ctactcgggaccttacggggacatcggttggagactgccagggaccatgtttggccat	3521
DB	2529	CTACTCGGACCTTATGGGACATGGTTTGGACAGTACCAGGGACACAGTTTACCCTAT	2588
QY	3522	tgaactatcttccaccccgagaagactgcctgactctgtgtgagatgaagctctctgggtg	3581
DB	2589	CGACTATTACTTCCACCCCGCAAGAACCTTGCCTGATCTGTGGAGATGAAGCTTCTGGTTG	2648
QY	3582	tcactatggagctctcacatgtggagctgcaagctcttcttcaaaagaccctgaagg	3641
DB	2649	TCACTACGGAGCTCTCACTGTGGCAGCTGCAAGGCTCTCTTCAAAAGAGCTCGCGAAGG	2708
QY	3642	gaaacagaagtactctgcccagcagaaaaatgtatggcactattgataaaattccgaaggaa	3701
DB	2709	GAACACAGAATATCTATGTGCCAGCAGAAATGATTGACACCTTGCATTAATTTTCGGAGAA	2768
QY	3702	aaattgtccactctgtcgtctctcggaatatgttatgaagcagggatgactctggagcccg	3761
DB	2769	AAATGTCCATCTGCTGCTCCGGAAATGTTTATGAAGCAGGGATGACTCTGGGAGCTCG	2828
QY	3762	gaagctgaagaagacttgtaactcgaactacagaggaaggaagcgttccaagccac	3821
DB	2829	TAGCTGAAGAAACTTGGAAATCTCAACTACAGAGAGAGGAGAAATCTCCAGTCTGTG	2888
QY	3822	cagccccactggagagacaacccgaagctgacagtgctcacacattgaaagctatgaattg	3881

```

|||||
2889 TAGCCCCACTGAGGACCCATCCCAAGAAGATGACTGTATCAACACATTGAAGGCTATGAATG 2948
QY 3882 teagcceafttttgaatgctcctggaagccattgagccaggtagtgtgtctggaaca 3941
Db 2949 TCAACCTATCTTTTAATGTCCTCGGAAGCCATTGAGCCAGGAGTGTGTGTCGCCGACA 3008
QY 3942 qgacaacaccagccgactcctttgcagccttgcctctctctctagccctcaatgaactgggaga 4001
Db 3009 TGACAACACCCAGCCTGATCTCTTGCTGCCTTGTATCTAGTCTCAACGAGCTTGGCGA 3068
QY 4002 gagacagcttgtaacacgtggtcaagtgggcgaagcccttgctcgcttccgaacttaca 4061
Db 3069 GAGACAGCTTGTCATATGTCGTAAGTGGGCCAAGGCCCTGGCTGGCTTCCCAACTTGCA 3128
QY 4062 cgtgcacacacagatggctgctcattccagctactcctcgtggtgggctcatggtgtttgcca 4121
Db 3129 TGTGGATGACCCAGATGGCAGTCAATTCAGTATCTCTGGATGGGACTGATGGTATTGGCAT 3188
QY 4122 gggctggcgatccttcaccaatgtcaactccaggatgctctactcgcgcctctgctctggt 4181
Db 3189 GGGTTGGCGGCTCTTCACTAATGTCAACTAGGATGCTACTTTGCACCTGACCTGGT 3248
QY 4182 ttccaatgagtcaccgcacatgcacaaagtcggagtgatcacgccagtgctgcgaatgagca 4241
Db 3249 TTTCATGAGTATCGCATGCACAAGTCTCGAATGTACAGCCAGTCGCTGAGATGAGCA 3308
QY 4242 cctctcgaaggtttggatggctcacaatcacccccccaggaattcctgtgcatagaagc 4301
Db 3309 CCTTCTCAAGAGTTGGATGGCTCCAGATAACCCCCCAGGAATTCCTGTGCATGAAGC 3368
QY 4302 actgctactcttoagcattctccagtgatgggtggtgaaatacaaaaattctttgatga 4361
Db 3369 ACTGCTACTCTTCAGCATTTATTCAGTGGTGGCTGGAATAATCAAAAATCTTTGTATGA 3428
QY 4362 acttccaatgaactacatacaagaaactcgtatcattcattgcatacgaagaaaaatcc 4421
Db 3429 ACTTCAATGAAGTACATCAAGAACTTGATCGCATTTGATGCATGAAGAAAAATCC 3488
QY 4422 cacactcgtcgaagcgtctctaccagctaccgaagctcctggactcgtgcagcctat 4481
Db 3489 CACATCTGCTCAAGCGGCTCTTACCAGCTCACAAAGCTCCTGGATTCCTGTCAGCCTAT 3548
QY 4482 tgcgagagactcatcagttcaacttttcaactgtcaactgatacaagtcacacatggtgagct 4541
Db 3549 TGCAAAGAGAGCTGCATCAATTCATTTGACCTGCTTAATCAAGTCCCATATGGTGAAGCT 3608
QY 4542 ggaacttccggaaatgatggcagagatcattctgtgcaagtgcccaagatccttcttgg 4601
Db 3609 GGACTTTCTGAAATGATGCGCAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGG 3668
QY 4602 gaaagtcaagcccatctatttcacacaccgaagtaagaacttggaacccatttccccacc 4661
Db 3669 GAAAGTCAAGCCATCATTTTCCACACACACAGATGAAGATTTGGAAACCCCTAATACCAAC 3728
QY 4662 ccaactcatgcccccttcagatgtctctgcctgt--tataactcgactactcctct 4719
Db 3729 CCACCTTGT--TCCCTTTTCAGATGCTCTCTGCGCTGTTATATATACTCTGCACTACTTCTCT 3787
QY 4720 gcagtgacct--ggggaatttctctatatgtgatgtacagctgtgcatgaacatgttctctgaa 4778
Db 3788 GCATTTGCTTGGGGAAATTCCTACTGATGTACAGTCTGTGATGAACATGTTCCCCAA 3847
QY 4779 ttctattgtgggcttttttttctctctctctctctctctctctctctctctctctcccta 4838
Db 3848 -----GTTCTATTTCTGGGCTTTTTCCTCTCTCTTTCTTTCTTTCTTTCTCTCTGCTC 3896
QY 4839 tetaacctcccatggcaccttcagactttgtcttcccatattgtggtcctcatctgtgtttt 4898
Db 3897 TTTTACCTTCCCATGGCACAATTTTGAATCCGCTCGGTGTTGTGGCTCCTCGCTGTGTTT 3956
QY 4899 gaatgggtgtgtatgccttttaaatctgtgatgtatcctcatatggccaggtgtaagttgt 4958
|||||

```

```

Db 3957 GAGTTTGTGTATTCTTCAAGTCTGTGATGATCTTCTTGTGGCCAGTGTCAA-CTGT 4015
QY 4959 gcttgttttacagcac--tactctgtgcagccacacaaacgtttacttatcttatccac 5016
Db 4016 GCTTGTGTTATAGCACTGTGCTGTGTGCCCAACCAAGCAAAATGTTTACTTACCTTATGCCAT 4075
QY 5017 ggaagtttagagagcctaagattatctggggaatcaaaacaaaaaacaagcaacaa 5074
Db 4076 GGCAAGTTTAGAGAGCTATAAGTATCTTGGAGAGAAACAACAGAGAGAGTAAAAAA 4133

RESULT 15
RATANDREC 4137 bp mRNA ROD 27-APR-1993
LOCUS Rat androgen receptor (AR) mRNA, complete cds.
DEFINITION M20133
ACCESSION M20133.1 GI:202895
VERSION androgen receptor.
KEYWORDS Rat, epididymal cdna to mRNA, clone rAREpl.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 4137)
AUTHORS Tan,J., Joseph,D.R., Quarmby,V.E., Lubahn,D.B., Sar,M., French,F.S.
and Wilson,E.M.
TITLE The rat androgen receptor: Primary structure, autoregulation of its
messenger ribonucleic acid, and immunocytochemical localization of
the receptor protein.
JOURNAL Mol. Endocrinol. 2, 1276-1285 (1988)
MEDLINE 89112209
COMMENT Draft entry and computer readable sequence [1] kindly submitted by
E.M.Wilson, 18-AUG-1988.
FEATURES
    source
        1..4137
            /organism="Rattus norvegicus"
            /db_xref="taxon:10116"
            994..3702
            /note="androgen receptor"
            /protein_id="AAA40733.1"
            /codon_start=1
            /db_xref="GI:202896"
            /translation="MEVQLGLGRVYPRPSKTYRGAFONLFQSVREAIQNGPRRPEA
            AS1APGACLOQROETSPRRRRQQHPEDSGPOAHIRGTGYLALBEEQPSQQQSAS
            EGHPECLPEPGAATAPKGLPQPPAPPDQDDSAAPSTLSLGLPTFLGLSSCSADI
            KDILSEAGTLMQLQQQQQQQQQQQQQQQVEISEGSSVRAREATGAPSSSK
            DSYLGNSTTSDSAKELCAVSYSMGLGVEALEHLSPGEOLGDCMYATSLGGPPAVR
            PTPCAPLAECKGLSLDEGPKGTEETAEYSFKGKAGLEGSGLGCSGSSEAGSSGT
            LEIPSSLSLYKSGAVDEAAAYONRDYVNFPLALSGPPHPPPHPHARIKLENPSDY
            SAWAAAAQCRYGLDLSLHGSGVAGSPSTGSPPTATSSWHTLTATBEGQLYGPGGYG
            SSPSDAPVAPYGYTRPQGLASQEGDFSASEVYVPGVNVNRYPIYPSPCVYKSEMG
            WMEYSGPYGDMRLDSTRDHLVLTIDYFPQKTLICGDEASGHYALTCGSKVFF
            KRAAEKGKYLACSRNDCTIDKFRKNCPSRLKCYEAGMTLCAARKKLKGLKLOE
            EGENSAGSPTEPSOKVTHIEGYECPFLNLVLEAIEPVVCAAGHNNQDPSFAA
            LLSNELGEROLVHVYKAKALPGFENLHVDDOMAVIOYSWMLVMVFMGWSFTNV
            NSRMLYFAPDLVFNRYMRHKSMTYSCVRRHLHSQBEFGNLQITPOEFLCMKALLFSI
            IPVDLKNKFFDELRMNLIKELDRITACKRNKNTSCSRKPYQLTKLSDVQPIAREL
            HQFTDLLIKSHMYSVDFFPENMAEIIISVQVPKILSGKVKPIYFHTQ"
BASE COUNT 979 a 1149 c 1086 g 923 t
ORIGIN

```

Query Match 48.7%; Score 2473.2; DB 10; Length 4137;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 3406; Conservative 0; Mismatches 666; Indels 308; Gaps 23;

QY 758 atcccgggagagccactgtcgtgagagcgggaacggtcccgagcaagccagagcgaga 817

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 1 ATCCCTTAGGAGCAGCCTCTCTGGAGAACACAG-AGGTCCGGAGCAAAACCTTGAGGCTGA 59

QY 818 ggagcgagcagagggaaaaagggcccnagcttagccgctccagtgctgtacnagncgcaa 877

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 60 GAGGGCATCAGAGGGGAAAAA-CTGAGTTAGCCACTCCAGTCCATACAGAGCTTAA 117
